=> FI HCA;D QUE L4;D BIB 1-
FILE 'HCA' ENTERED AT 11:28:13 ON 14 MAY 96
USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT

FILE COVERS 1967 - 8 May 1996 (960508/ED) VOL 124 ISS 20

To help control your online searching costs, consider using the HCA File when using the FSEARCH command or when conducting SmartSELECT searches with large numbers of terms.

Thesauri are now available for the WIPO International Patent Classifications (IPC) editions 1-6 in the /IC1, /IC2, /IC3, /IC4, /IC5, and /IC (/IC6) fields, respectively. The thesauri in the /IC5 and /IC fields also include the corresponding catchword terms from the IPC subject headings and subheadings.

```
L1            40 SEA FILE=REGISTRY GAAGTTCCTATTC/SQSN
L2            34 SEA FILE=REGISTRY GTATAGGAACTTC/SQSN
L3            29 SEA FILE=REGISTRY L1(L)L2
L4             5 SEA FILE=HCA L3
```

L4 ANSWER 1 OF 5 HCA COPYRIGHT 1996 ACS
AN 123:331869 HCA
TI The role of DNA bending in Flp-mediated site-specific recombination
AU Luetke, Karen H.; Sadowski, Paul D.
CS Dep. Molecular and Medical Genetics, Univ. Toronto, Toronto, ON, M5S 1A8, Can.
SO J. Mol. Biol. (1995), 251(4), 493-506
CODEN: JMOBAK; ISSN: 0022-2836
DT Journal
LA English

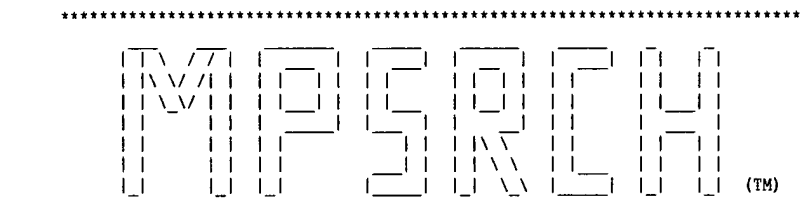
L4 ANSWER 2 OF 5 HCA COPYRIGHT 1996 ACS
AN 121:126321 HCA
TI In vivo excision and amplification of large segments of the Escherichia coli genome
AU Posfai, Gyorgy; Koob, Michael; Hradecna, Zdenka; Hasan, Noaman; Filutowicz, Marcin; Szybalski, Waclaw
CS McArdle Lab. Cancer Res., Univ. Wisconsin, Madison, WI, 53706, USA
SO Nucleic Acids Res. (1994), 22(12), 2392-8
CODEN: NARHAD; ISSN: 0305-1048
DT Journal
LA English

L4 ANSWER 3 OF 5 HCA COPYRIGHT 1996 ACS
AN 119:198292 HCA
TI Ligation of synthetic activated DNA substrates by site-specific recombinases and topoisomerase I

AU ' Guohua; Luetke, Karen; Juby, Carl D.; Brousseau, Roland; Sadowski, Paul
CS Dep. Mol. Med. Genet., Univ. Toronto, Toronto, ON, M5S 1A8, Can.
SO J. Biol. Chem. (1993), 268(5), 3683-9
CODEN: JBCHA3; ISSN: 0021-9258
DT Journal
LA English

L4 ANSWER 4 OF 5 HCA COPYRIGHT 1996 ACS
AN 114:1432 HCA
TI Nucleotide sequence of a gene which enhances the activity of glyoxalase I in Saccharomyces cerevisiae
AU Inoue, Yoshiharu; Feng, Ling; Bong-Young, Choi; Ginya, Harumi; Murata, Kousaka; Kimura, Akira
CS Res. Inst. Food Sci., Kyoto Univ., Uji, 611, Japan
SO Biotechnol. Appl. Biochem. (1990), 12(3), 341-5
CODEN: BABIEC; ISSN: 0885-4513
DT Journal
LA English

L4 ANSWER 5 OF 5 HCA COPYRIGHT 1996 ACS
AN 94:12618 HCA
TI Nucleotide sequence of the yeast plasmid
AU Hartley, James L.; Donelson, John E.
CS Dep. Biochem., Univ. Iowa, Iowa City, IA, 52242, USA
SO Nature (London) (1980), 286(5776), 860-5
CODEN: NATUAS; ISSN: 0028-0836
DT Journal
LA English



Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue May 14 11:40:28 1996; MasPar time 2.46 Seconds
254.769 Million cell updates/sec

Tabular output not generated.

Title: >FLP
Description: (1-34) from frt.seq
Perfect Score: 26
N.A. Sequence: 1 gaagttcctattcNNNNNNNNgtataggaacttc 34
Comp: cttcaaggataagNNNNNNNNcatatccttgaag

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 30947 seqs, 9219370 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:51 2:52 3:53 4:54 5:55 6:PCT90 7:PCT91 8:PCT92 9:PCT93
10:PCT94 11:PCT95

Statistics: Mean 5.064; Variance 2.619; scale 1.934

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	26	100.0	34	8	PCT-US92-0	Sequence 3, Applicatio	4.35e-08
	2	26	100.0	68	8	PCT-US92-0	Sequence 4, Applicatio	4.35e-08
	3	26	100.0	7859	4	US-07-854-	Sequence 4, Applicatio	4.35e-08
c	4	22	84.6	34	8	PCT-US92-0	Sequence 3, Applicatio	2.38e-05
c	5	22	84.6	68	8	PCT-US92-0	Sequence 4, Applicatio	2.38e-05
c	6	22	84.6	7859	4	US-07-854-	Sequence 4, Applicatio	2.38e-05
c	7	14	53.8	1947	9	PCT-US93-0	Sequence 4, Applicatio	2.71e+00
c	8	14	53.8	1997	9	PCT-US93-0	Sequence 6, Applicatio	2.71e+00
c	9	13	50.0	1065	4	US-08-292-	Sequence 5, Applicatio	1.01e+01
c	10	13	50.0	1689	4	US-07-991-	Sequence 41, Applicati	1.01e+01

c	11	13	50.0	1947	4	US-07-998-	Sequence 2, Applicatio	1.01e+01
c	12	13	50.0	1947	5	US-08-463-	Sequence 2, Applicatio	1.01e+01
c	13	13	50.0	1947	8	PCT-US92-1	Sequence 2, Applicatio	1.01e+01
c	14	13	50.0	1947	11	PCT-US95-0	Sequence 24, Applicati	1.01e+01
c	15	13	50.0	1988	4	US-07-750-	Sequence 15, Applicati	1.01e+01
c	16	13	50.0	2188	3	US-07-882-	Sequence 4, Applicatio	1.01e+01
c	17	13	50.0	2553	11	PCT-US95-0	Sequence 5, Applicatio	1.01e+01
c	18	13	50.0	6560	9	PCT-US93-0	Sequence 1, Applicatio	1.01e+01
	19	13	50.0	7493	10	PCT-US94-1	Sequence 5, Applicatio	1.01e+01
c	20	13	50.0	7493	10	PCT-US94-1	Sequence 5, Applicatio	1.01e+01
	21	13	50.0	7653	9	PCT-US93-0	Sequence 1, Applicatio	1.01e+01
	22	13	50.0	7653	9	PCT-US93-0	Sequence 18, Applicati	1.01e+01
c	23	13	50.0	8316	3	US-07-753-	Sequence 4, Applicatio	1.01e+01
c	24	13	50.0	8457	4	US-07-991-	Sequence 1, Applicatio	1.01e+01
c	25	13	50.0	9115	3	US-07-753-	Sequence 3, Applicatio	1.01e+01
	26	12	46.2	31	4	US-08-278-	Sequence 4, Applicatio	3.58e+01
c	27	12	46.2	237	9	PCT-US93-1	Sequence 9, Applicatio	3.58e+01
c	28	12	46.2	360	4	US-08-026-	Sequence 3, Applicatio	3.58e+01
c	29	12	46.2	1229	10	PCT-US94-0	Sequence 54, Applicati	3.58e+01
	30	12	46.2	1635	3	US-08-133-	Sequence 3, Applicatio	3.58e+01
c	31	12	46.2	1749	4	US-08-277-	Sequence 2, Applicatio	3.58e+01
c	32	12	46.2	2230	4	US-08-217-	Sequence 5, Applicatio	3.58e+01
	33	12	46.2	2339	11	PCT-US95-0	Sequence 1, Applicatio	3.58e+01
c	34	12	46.2	2356	4	US-08-105-	Sequence 222, Applicat	3.58e+01
c	35	12	46.2	2679	7	PCT-US91-0	Sequence 11, Applicati	3.58e+01
c	36	12	46.2	2757	9	PCT-US93-0	Sequence 79, Applicati	3.58e+01
	37	12	46.2	3033	11	PCT-US95-0	Sequence 111, Applicat	3.58e+01
c	38	12	46.2	3033	11	PCT-US95-0	Sequence 111, Applicat	3.58e+01
	39	12	46.2	3513	8	PCT-US92-1	Sequence 5, Applicatio	3.58e+01
	40	12	46.2	3513	4	US-08-278-	Sequence 1, Applicatio	3.58e+01
c	41	12	46.2	4131	8	PCT-US92-0	Sequence 38, Applicati	3.58e+01
c	42	12	46.2	4131	3	US-07-872-	Sequence 38, Applicati	3.58e+01
c	43	12	46.2	4146	11	PCT-US95-0	Sequence 15, Applicati	3.58e+01
c	44	12	46.2	5261	3	US-08-045-	Sequence 3, Applicatio	3.58e+01
c	45	12	46.2	8298	9	PCT-US93-0	Sequence 1, Applicatio	3.58e+01

ALIGNMENTS

RESULT 1
ID PCT-US92-01899-3 STANDARD; DNA; UNC; 34 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 3, Application PC/TUS9201899.
CC Sequence 3, Application PC/TUS9201899
CC GENERAL INFORMATION:
CC APPLICANT: WAHL, DR., GEOFFREY M.
CC APPLICANT: O'GORMAN DR., STEPHEN V.
CC TITLE OF INVENTION: FLP-MEDIATED GENE MODIFICATION IN
CC TITLE OF INVENTION: MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL
CC TITLE OF INVENTION: THEREFOR
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
CC STREET: 444 South Flower Street, Suite 2000
CC CITY: Los Angeles
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 90071
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25

```
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/01899
CC        FILING DATE:  19920306
CC        CLASSIFICATION:  800
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  REITER MR., STEPHEN E.
CC        REGISTRATION NUMBER:  31192
CC        REFERENCE/DOCKET NUMBER:  P31 8929
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (619) 535-9001
CC        TELEFAX:  (619) 535-8949
CC    INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  34 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  unknown
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  cDNA
CC      ORIGINAL SOURCE:
CC        INDIVIDUAL ISOLATE:  FLP recombination target site
SQ   Sequence 34 BP; 11 A; 6 C; 6 G; 11 T; 0 other;

  Query Match          100.0%;  Score 26;  DB 8;  Length 34;
  Best Local Similarity  76.5%;  Pred. No. 4.35e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db      1 GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC 34
          ||||||||||||||        ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    2
ID   PCT-US92-01899-4 STANDARD; DNA; UNC; 68 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application PC/TUS9201899.
CC   Sequence 4, Application PC/TUS9201899
CC    GENERAL INFORMATION:
CC      APPLICANT:  WAHL, DR., GEOFFREY M.
CC      APPLICANT:  O'GORMAN DR., STEPHEN V.
CC      TITLE OF INVENTION:  FLP-MEDIATED GENE MODIFICATION IN
CC      TITLE OF INVENTION:  MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL
CC      TITLE OF INVENTION:  THEREFOR
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
CC        STREET:  444 South Flower Street, Suite 2000
CC        CITY:  Los Angeles
CC        STATE:  CA
CC        COUNTRY:  USA
CC        ZIP:  90071
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/01899
CC        FILING DATE:  19920306
CC        CLASSIFICATION:  800
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  REITER MR., STEPHEN E.
CC        REGISTRATION NUMBER:  31192
CC        REFERENCE/DOCKET NUMBER:  P31 8929
```



```
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (619) 535-9001
CC        TELEFAX:  (619) 535-8949
CC    INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  68 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  unknown
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  cDNA
CC      ORIGINAL SOURCE:
CC        INDIVIDUAL ISOLATE:  Synthetic oligonucleotide
SQ   Sequence 68 BP; 19 A; 16 C; 14 G; 19 T; 0 other;

  Query Match          100.0%;  Score 26;  DB 8;  Length 68;
  Best Local Similarity  76.5%;  Pred. No. 4.35e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     34 GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC 67
          ||||||||||||||        ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    3
ID   US-07-854-596B-4 STANDARD; DNA; UNC; 7859 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application US/07854596B.
CC   Sequence 4, Application US/07854596B
CC   Patent No. 5434073
CC    GENERAL INFORMATION:
CC      APPLICANT:  Dawson, Keith M
CC      APPLICANT:  Hunter, Michael G
CC      APPLICANT:  Czaplewski, Lloyd G
CC      TITLE OF INVENTION:  Proteins and nucleic acids
CC      NUMBER OF SEQUENCES:  73
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Dr. John J. McDonnell
CC        STREET:  Ten South Wacker Drive, Suite 3000
CC        CITY:  Chicago
CC        STATE:  IL
CC        COUNTRY:  USA
CC        ZIP:  60606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/854,596B
CC        FILING DATE:  03-JUN-1992
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  McDonnell, John J
CC        REGISTRATION NUMBER:  26,949
CC        REFERENCE/DOCKET NUMBER:  92,337
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  312-715-1000
CC        TELEFAX:  312-715-1234
CC        TELEX:  910-221-5317
CC    INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7859 base pairs
CC        TYPE:  nucleic acid
```

```
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  circular
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  1..7859
CC        OTHER INFORMATION:  /note= "sequence of plasmid pSW6"
SQ   Sequence 7859 BP; 2317 A; 1656 C; 1600 G; 2286 T; 0 other;

  Query Match          100.0%;  Score 26;  DB 4;  Length 7859;
  Best Local Similarity  76.5%;  Pred. No. 4.35e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC 3164
           ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    4
ID   PCT-US92-01899-3 STANDARD; DNA; UNC; 34 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application PC/TUS9201899.
CC   Sequence 3, Application PC/TUS9201899
CC    GENERAL INFORMATION:
CC      APPLICANT:  WAHL, DR., GEOFFREY M.
CC      APPLICANT:  O'GORMAN DR., STEPHEN V.
CC      TITLE OF INVENTION:  FLP-MEDIATED GENE MODIFICATION IN
CC      TITLE OF INVENTION:  MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL
CC      TITLE OF INVENTION:  THEREFOR
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
CC        STREET:  444 South Flower Street, Suite 2000
CC        CITY:  Los Angeles
CC        STATE:  CA
CC        COUNTRY:  USA
CC        ZIP:  90071
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/01899
CC        FILING DATE:  19920306
CC        CLASSIFICATION:  800
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  REITER MR., STEPHEN E.
CC        REGISTRATION NUMBER:  31192
CC        REFERENCE/DOCKET NUMBER:  P31 8929
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (619) 535-9001
CC        TELEFAX:  (619) 535-8949
CC     INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  34 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  unknown
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  cDNA
CC      ORIGINAL SOURCE:
CC        INDIVIDUAL ISOLATE:  FLP recombination target site
SQ   Sequence 34 BP; 11 A; 6 C; 6 G; 11 T; 0 other;
```



```
  Query Match          84.6%;  Score 22;  DB 8;  Length 34;
  Best Local Similarity  70.6%;  Pred. No. 2.38e-05;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db      1 GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC 34
          |||||||||||| |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT    5
ID   PCT-US92-01899-4 STANDARD; DNA; UNC; 68 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application PC/TUS9201899.
CC   Sequence 4, Application PC/TUS9201899
CC    GENERAL INFORMATION:
CC      APPLICANT:  WAHL, DR., GEOFFREY M.
CC      APPLICANT:  O'GORMAN DR., STEPHEN V.
CC      TITLE OF INVENTION:  FLP-MEDIATED GENE MODIFICATION IN
CC      TITLE OF INVENTION:  MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL
CC      TITLE OF INVENTION:  THEREFOR
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
CC        STREET:  444 South Flower Street, Suite 2000
CC        CITY:  Los Angeles
CC        STATE:  CA
CC        COUNTRY:  USA
CC        ZIP:  90071
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/01899
CC        FILING DATE:  19920306
CC        CLASSIFICATION:  800
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  REITER MR., STEPHEN E.
CC        REGISTRATION NUMBER:  31192
CC        REFERENCE/DOCKET NUMBER:  P31 8929
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (619) 535-9001
CC        TELEFAX:  (619) 535-8949
CC     INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  68 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  unknown
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  cDNA
CC      ORIGINAL SOURCE:
CC        INDIVIDUAL ISOLATE:  Synthetic oligonucleotide
SQ   Sequence 68 BP; 19 A; 16 C; 14 G; 19 T; 0 other;

  Query Match          84.6%;  Score 22;  DB 8;  Length 68;
  Best Local Similarity  70.6%;  Pred. No. 2.38e-05;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db     34 GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC 67
          |||||||||||| |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1
```

```
RESULT    6
ID   US-07-854-596B-4 STANDARD; DNA; UNC; 7859 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application US/07854596B.
CC   Sequence 4, Application US/07854596B
CC   Patent No. 5434073
CC    GENERAL INFORMATION:
CC      APPLICANT:  Dawson, Keith M
CC      APPLICANT:  Hunter, Michael G
CC      APPLICANT:  Czaplewski, Lloyd G
CC      TITLE OF INVENTION:  Proteins and nucleic acids
CC      NUMBER OF SEQUENCES:  73
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Dr. John J. McDonnell
CC        STREET:  Ten South Wacker Drive, Suite 3000
CC        CITY:  Chicago
CC        STATE:  IL
CC        COUNTRY:  USA
CC        ZIP:  60606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/854,596B
CC        FILING DATE:  03-JUN-1992
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  McDonnell, John J
CC        REGISTRATION NUMBER:  26,949
CC        REFERENCE/DOCKET NUMBER:  92,337
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  312-715-1000
CC        TELEFAX:  312-715-1234
CC        TELEX:  910-221-5317
CC     INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7859 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  circular
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  1..7859
CC        OTHER INFORMATION:  /note= "sequence of plasmid pSW6"
SQ   Sequence 7859 BP; 2317 A; 1656 C; 1600 G; 2286 T; 0 other;

  Query Match           84.6%;  Score 22;  DB 4;  Length 7859;
  Best Local Similarity  70.6%;  Pred. No. 2.38e-05;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   3131 GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC 3164
          |||||||||||| |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT    7
ID   PCT-US93-01720-4 STANDARD; DNA; UNC; 1947 BP.
AC   xxxxxx
```



```
DT   01-JAN-1900
DE   Sequence 4, Application PC/TUS9301720.
CC   Sequence 4, Application PC/TUS9301720
CC    GENERAL INFORMATION:
CC      APPLICANT:  Renauld, Jean-Christophe
CC      APPLICANT:  Druez, Catherine
CC      APPLICANT:  Van Snick, Jacques
CC      TITLE OF INVENTION:  Nucleic Acid Sequences Coding For
CC      TITLE OF INVENTION:  Or Complementary To Nucleic Acid Sequences Coding
  For
CC      TITLE OF INVENTION:  Interleukin 9 Receptor
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Felfe & Lynch
CC        STREET:  805 Third Avenue
CC        CITY:  New York City
CC        STATE:  New York
CC        COUNTRY:  USA
CC        ZIP:  10022
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC        COMPUTER:  IBM PS/2
CC        OPERATING SYSTEM:  PC-DOS
CC        SOFTWARE:  Wordperfect
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/01720
CC        FILING DATE:  19930225
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/847,347
CC        FILING DATE:  09-MARCH-1992
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hanson, Norman D.
CC        REGISTRATION NUMBER:  30,946
CC        REFERENCE/DOCKET NUMBER:  LUD 264
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 688-9200
CC        TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1947 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
SQ   Sequence 1947 BP; 379 A; 582 C; 586 G; 400 T; 0 other;

  Query Match           53.8%;  Score 14;  DB 9;  Length 1947;
  Best Local Similarity  60.7%;  Pred. No. 2.71e+00;
  Matches    17;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   1121 CCTCTACAGTGTACACAATGGGAACTTC 1148
          ||| |||         ||| ||||||||
Cp     28 cctatacNNNNNNNNgaataggaacttc 1


RESULT    8
ID   PCT-US93-01720-6 STANDARD; DNA; UNC; 1997 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 6, Application PC/TUS9301720.
CC   Sequence 6, Application PC/TUS9301720
CC    GENERAL INFORMATION:
CC      APPLICANT:  Renauld, Jean-Christophe
CC      APPLICANT:  Druez, Catherine
```

```
CC      APPLICANT:  Van Snick, Jacques
CC      TITLE OF INVENTION:  Nucleic Acid Sequences Coding For
CC      TITLE OF INVENTION:  Or Complementary To Nucleic Acid Sequences Coding
  For
CC      TITLE OF INVENTION:  Interleukin 9 Receptor
CC      NUMBER OF SEQUENCES:  6
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Felfe & Lynch
CC        STREET:  805 Third Avenue
CC        CITY:  New York City
CC        STATE:  New York
CC        COUNTRY:  USA
CC        ZIP:  10022
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC        COMPUTER:  IBM PS/2
CC        OPERATING SYSTEM:  PC-DOS
CC        SOFTWARE:  Wordperfect
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/01720
CC        FILING DATE:  19930225
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/847,347
CC        FILING DATE:  09-MARCH-1992
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hanson, Norman D.
CC        REGISTRATION NUMBER:  30,946
CC        REFERENCE/DOCKET NUMBER:  LUD 264
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (212) 688-9200
CC        TELEFAX:  (212) 838-3884
CC    INFORMATION FOR SEQ ID NO:  6:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1997 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
SQ   Sequence 1997 BP; 388 A; 612 C; 593 G; 404 T; 0 other;

  Query Match           53.8%;  Score 14;  DB 9;  Length 1997;
  Best Local Similarity  60.7%;  Pred. No. 2.71e+00;
  Matches    17;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   1171 CCTCTACAGTGTACACAATGGGAACTTC 1198
            ||| |||         ||| ||||||||
Cp     28 cctatacNNNNNNNNgaataggaacttc 1


RESULT    9
ID   US-08-292-549-5 STANDARD; DNA; UNC; 1065 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application US/08292549.
CC   Sequence 5, Application US/08292549
CC   Patent No. 5464938
CC    GENERAL INFORMATION:
CC      APPLICANT:  Smith, Craig A.
CC      APPLICANT:  Goodwin, Raymond G.
CC      TITLE OF INVENTION:  Isolated Viral Protein TNF Antagonists
CC      NUMBER OF SEQUENCES:  10
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Immunex Corporation
CC        STREET:  51 University Street
```



```
CC        CITY:  Seattle
CC        STATE:  Washington
CC        COUNTRY:  USA
CC        ZIP:  98101
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/292,549
CC        FILING DATE:
CC        CLASSIFICATION:  530
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  07/963,330
CC        FILING DATE:  10/19/92
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Perkins, Patricia A.
CC        REGISTRATION NUMBER:  34,693
CC        REFERENCE/DOCKET NUMBER:  2602-A
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (206) 587-0430
CC        TELEFAX:  (206) 233-0644
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1065 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Cowpox virus
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  1..1065
SQ   Sequence 1065 BP; 327 A; 248 C; 197 G; 293 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 4;  Length 1065;
  Best Local Similarity  60.0%;  Pred. No. 1.01e+01;
  Matches    15;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    384 TTCCCAAACAAAGTGTGGAATAGGA 408
          |||| | ||        ||||||||
Cp     30 ttcctatacNNNNNNNNgaatagga 6


RESULT   10
ID   US-07-991-867B-41 STANDARD; DNA; UNC; 1689 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 41, Application US/07991867B.
CC   Sequence 41, Application US/07991867B
CC   Patent No. 5476781
CC    GENERAL INFORMATION:
CC      APPLICANT:  Moyer, Richard W.
CC      APPLICANT:  Hall, Richard L.
CC      APPLICANT:  Gruidl, Michael E.
CC      TITLE OF INVENTION:  No. 5476781el Entomopoxvirus Expression System
CC      NUMBER OF SEQUENCES:  66
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  David R. Saliwanchik
CC        STREET:  2421 N.W. 41st Street, Suite A-1
```

```
CC        CITY:  Gainesville
CC        STATE:  FL
CC        COUNTRY:  USA
CC        ZIP:  32606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/991,867B
CC        FILING DATE:  12-DEC-1992
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  WO 92/14818
CC        FILING DATE:  12-FEB-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/827,685
CC        FILING DATE:  30-JAN-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/657,584
CC        FILING DATE:  19-FEB-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Saliwanchik, David R.
CC        REGISTRATION NUMBER:  31,794
CC        REFERENCE/DOCKET NUMBER:  UF114.C3
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  904-375-8100
CC        TELEFAX:  904-372-5800
CC    INFORMATION FOR SEQ ID NO:  41:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1689 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 1689 BP; 595 A; 200 C; 149 G; 745 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 4;  Length 1689;
  Best Local Similarity  60.0%;  Pred. No. 1.01e+01;
  Matches    15;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    312 AAGTTTCTATATATTTACACGAATA 336
          ||||| |||||         |||||
Cp     33 aagttcctatacNNNNNNNNgaata 9


RESULT   11
ID   US-07-998-972A-2 STANDARD; DNA; UNC; 1947 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 2, Application US/07998972A.
CC   Sequence 2, Application US/07998972A
CC   Patent No. 5476777
CC    GENERAL INFORMATION:
CC      APPLICANT:  Holly, Richard D.
CC      APPLICANT:  Foster, Donald C.
CC      TITLE OF INVENTION:  METHODS FOR PRODUCING THROMBIN
CC      NUMBER OF SEQUENCES:  48
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Townsend and Townsend
CC        STREET:  One Market Plaza, Stewart Street Tower,
CC        STREET:  Twentieth Floor
CC        CITY:  San Francisco
```



```
CC        STATE:  CA
CC        COUNTRY:  USA
CC        ZIP:  94105
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/998,972A
CC        FILING DATE:  19921230
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/860,701
CC        FILING DATE:  31-MAR-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/816,281
CC        FILING DATE:  31-DEC-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Parmelee, Steven W
CC        REGISTRATION NUMBER:  31,990
CC        REFERENCE/DOCKET NUMBER:  13952-12-2
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  206-467-9600
CC        TELEFAX:  415-543-5043
CC    INFORMATION FOR SEQ ID NO:  2:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1947 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Homo sapiens
CC        TISSUE TYPE:  Hepatic
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  3..1847
SQ   Sequence 1947 BP; 439 A; 522 C; 609 G; 377 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 4;  Length 1947;
  Best Local Similarity  58.6%;  Pred. No. 1.01e+01;
  Matches    17;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   1204 TCCTGTACCCGCCCTGGGACAAGAACTTC 1232
          |||| |||        | | | |||||||
Cp     29 tcctatacNNNNNNNNgaataggaacttc 1


RESULT   12
ID   US-08-463-953-2 STANDARD; DNA; UNC; 1947 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 2, Application US/08463953.
CC   Sequence 2, Application US/08463953
CC   Patent No. 5502034
CC    GENERAL INFORMATION:
CC      APPLICANT:  Holly, Richard D.
CC      APPLICANT:  Foster, Donald C.
CC      TITLE OF INVENTION:  METHODS FOR PRODUCING THROMBIN
CC      NUMBER OF SEQUENCES:  48
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Townsend and Townsend
CC        STREET:  One Market Plaza, Stewart Street Tower,
CC        STREET:  Twentieth Floor
```

```
CC        CITY:  San Francisco
CC        STATE:  CA
CC        COUNTRY:  USA
CC        ZIP:  94105
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/463,953
CC        FILING DATE:
CC        CLASSIFICATION:  514
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/860,701
CC        FILING DATE:  31-MAR-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/816,281
CC        FILING DATE:  31-DEC-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Parmelee, Steven W
CC        REGISTRATION NUMBER:  31,990
CC        REFERENCE/DOCKET NUMBER:  13952-12-2
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  206-467-9600
CC        TELEFAX:  415-543-5043
CC    INFORMATION FOR SEQ ID NO:  2:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1947 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Homo sapiens
CC        TISSUE TYPE:  Hepatic
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  3..1847
SQ   Sequence 1947 BP; 439 A; 522 C; 609 G; 377 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 5;  Length 1947;
  Best Local Similarity  58.6%;  Pred. No. 1.01e+01;
  Matches   17;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   1204 TCCTGTACCCGCCCTGGGACAAGAACTTC 1232
           |||| |||        | | | |||||||
Cp     29 tcctatacNNNNNNNNgaataggaacttc 1


RESULT   13
ID   PCT-US92-11357-2 STANDARD; DNA; UNC; 1947 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 2, Application PC/TUS9211357.
CC   Sequence 2, Application PC/TUS9211357
CC    GENERAL INFORMATION:
CC      APPLICANT:  Holly, Richard D.
CC      APPLICANT:  Foster, Donald C.
CC      TITLE OF INVENTION:  METHODS FOR PRODUCING THROMBIN
CC      NUMBER OF SEQUENCES:  48
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Townsend and Townsend
CC        STREET:  One Market Plaza, Stewart Street Tower,
CC        STREET:  Twentieth Floor
```



```
CC        CITY:  San Francisco
CC        STATE:  CA
CC        COUNTRY:  USA
CC        ZIP:  94105
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/11357
CC        FILING DATE:  19921230
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/860,701
CC        FILING DATE:  31-MAR-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/816,281
CC        FILING DATE:  31-DEC-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Parmelee, Steven W
CC        REGISTRATION NUMBER:  31,990
CC        REFERENCE/DOCKET NUMBER:  13952-12-2
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  206-467-9600
CC        TELEFAX:  415-543-5043
CC    INFORMATION FOR SEQ ID NO:  2:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1947 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Homo sapiens
CC        TISSUE TYPE:  Hepatic
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  3..1847
SQ   Sequence 1947 BP; 439 A; 522 C; 609 G; 377 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 8;  Length 1947;
  Best Local Similarity  58.6%;  Pred. No. 1.01e+01;
  Matches   17;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   1204 TCCTGTACCCGCCCTGGGACAAGAACTTC 1232
           |||| |||        | | | |||||||
Cp     29 tcctatacNNNNNNNNgaataggaacttc 1


RESULT   14
ID   PCT-US95-07439-24 STANDARD; DNA; UNC; 1947 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 24, Application PC/TUS9507439.
CC   Sequence 24, Application PC/TUS9507439
CC    GENERAL INFORMATION:
CC      APPLICANT:
CC      APPLICANT:  NAME:     BOARD OF REGENTS, THE UNIVERSITY OF
CC      APPLICANT:  TEXAS SYSTEM
CC      APPLICANT:  STREET:   201 West 7th Street
CC      APPLICANT:  CITY:     Austin
CC      APPLICANT:  STATE:    Texas
CC      APPLICANT:  COUNTRY:  United States of America
CC      APPLICANT:  POSTAL CODE:  78701
```

```
CC      APPLICANT:  TELEPHONE NO: (512)499-4462
CC      APPLICANT:  TELEFAX:  (512)499-4523
CC      APPLICANT:  NAME:     THE SCRIPPS RESEARCH INSTITUTE
CC      APPLICANT:  STREET:   10666 North Torrey Pines Road
CC      APPLICANT:  CITY:     LaJolla
CC      APPLICANT:  STATE:    California
CC      APPLICANT:  COUNTRY:  United States of America
CC      APPLICANT:  POSTAL CODE:  92037
CC      TITLE OF INVENTION:  METHODS AND COMPOSITIONS
CC      TITLE OF INVENTION:  FOR THE SPECIFIC
CC      TITLE OF INVENTION:  COAGULATION OF VASCULATURE
CC      NUMBER OF SEQUENCES:  27
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Arnold, White & Durkee
CC        STREET:  P. O. Box 4433
CC        CITY:  Houston
CC        STATE:  Texas
CC        COUNTRY:  USA
CC        ZIP:  77210
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS, ASCII
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US95/07439
CC        FILING DATE:  Concurrently herewith
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 08/273,567
CC        FILING DATE:  11-JUN-1994
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  PARKER, DAVID L.
CC        REGISTRATION NUMBER:  32,165
CC        REFERENCE/DOCKET NUMBER:  UTFD433P--
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (512) 418-3000
CC        TELEFAX:  (713) 789-2679
CC        TELEX:  79-0924
CC    INFORMATION FOR SEQ ID NO:  24:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1947 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
SQ   Sequence 1947 BP; 439 A; 522 C; 609 G; 377 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 11;  Length 1947;
  Best Local Similarity  58.6%;  Pred. No. 1.01e+01;
  Matches    17;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   1204 TCCTGTACCCGCCCTGGGACAAGAACTTC 1232
           |||| |||        | | | |||||||
Cp     29 tcctatacNNNNNNNNgaataggaacttc 1


RESULT   15
ID   US-07-750-080A-15 STANDARD; DNA; UNC; 1988 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 15, Application US/07750080A.
CC   Sequence 15, Application US/07750080A
CC   Patent No. 5445953
CC    GENERAL INFORMATION:
CC      APPLICANT:  DORNER, F.
```



```
CC      APPLICANT:  SCHEIFLINGER, F.
CC      APPLICANT:  FALKNER, F. G.
CC      TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
CC      TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
CC      NUMBER OF SEQUENCES:  42
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Foley & Lardner
CC        STREET:  1800 Diagonal Road, Suite 500
CC        CITY:  Alexandria
CC        STATE:  VA
CC        COUNTRY:  USA
CC        ZIP:  22313-0299
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/750,080A
CC        FILING DATE:  19910826
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  BENT, Stephen A.
CC        REGISTRATION NUMBER:  29,768
CC        REFERENCE/DOCKET NUMBER:  30472/106 IMMU
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (703)836-9300
CC        TELEFAX:  (703)683-4109
CC        TELEX:  899149
CC    INFORMATION FOR SEQ ID NO:  15:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1988 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      IMMEDIATE SOURCE:
CC        CLONE:  pA1S1-PT (Fig. 5.1)
SQ   Sequence 1988 BP; 451 A; 529 C; 617 G; 391 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 4;  Length 1988;
  Best Local Similarity  58.6%;  Pred. No. 1.01e+01;
  Matches    17;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   1247 TCCTGTACCCGCCCTGGGACAAGAACTTC 1275
           |||| |||        | | | |||||||
Cp     29 tcctatacNNNNNNNNgaataggaacttc 1


RESULT   16
ID   US-07-882-925A-4 STANDARD; DNA; UNC; 2188 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application US/07882925A.
CC   Sequence 4, Application US/07882925A
CC   Patent No. 5315000
CC    GENERAL INFORMATION:
CC      APPLICANT:  Degen, Sandra J. F.
CC      TITLE OF INVENTION:  Gene for a growth factor and its cDNA and
CC      TITLE OF INVENTION:  protein
CC      NUMBER OF SEQUENCES:  7
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Gregory Lunn
CC        STREET:  Wood, Herron & Evans, 2700 Carew Tower
```

```
CC        CITY:  Cincinnati
CC        STATE:  Ohio
CC        COUNTRY:  USA
CC        ZIP:  45202
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Diskette, 3.50 inch,  800 Kb
CC        COMPUTER:  Apple Macintosh
CC        OPERATING SYSTEM:  Macintosh 6.0.3
CC        SOFTWARE:  Microsoft Word 4.0
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/882,925A
CC        FILING DATE:  19920514
CC        CLASSIFICATION:  530
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Lunn, Gregory
CC        REGISTRATION NUMBER:  29,945
CC        REFERENCE/DOCKET NUMBER:  CMC 57
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (513) 241-2324
CC        TELEFAX:  (513) 421-7269
CC    INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  2188 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA to mRNA
CC      ANTI-SENSE:  no
CC      ORIGINAL SOURCE:
CC        ORGANISM:  mouse
CC        STRAIN:  C57BL/6
CC        DEVELOPMENTAL STAGE:  adult
CC        TISSUE TYPE:  liver
CC      IMMEDIATE SOURCE:
CC        LIBRARY:  cDNA
CC        CLONE:  ML5-2
CC      POSITION IN GENOME:
CC        CHROMOSOME/SEGMENT:  mouse 9, Hgfl locus
CC        MAP POSITION:  Trf-Gnai-2-Hgfl-Cck
CC      FEATURE:
CC        IDENTIFICATION METHOD:  experimental
CC      PUBLICATION INFORMATION:
CC        RELEVANT RESIDUES IN SEQ ID NO:  4: 1 TO 2188
SQ   Sequence 2188 BP; 509 A; 608 C; 627 G; 444 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 3;  Length 2188;
  Best Local Similarity  60.9%;  Pred. No. 1.01e+01;
  Matches    14;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    641 AAGTTCCTAGACAAAGATCTGAA 663
          ||||||||| ||       |||
Cp     33 aagttcctatacNNNNNNNgaa 11


RESULT   17
ID   PCT-US95-07295-5 STANDARD; DNA; UNC; 2553 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application PC/TUS9507295.
CC   Sequence 5, Application PC/TUS9507295
CC    GENERAL INFORMATION:
CC      APPLICANT:  ALVES, KENNETH
CC      APPLICANT:  GUPTA, SUNIL K.
CC      APPLICANT:  HOLLIS, GREGORY F.
```



```
CC      TITLE OF INVENTION:  CONTRACEPTIVE VACCINE
CC      NUMBER OF SEQUENCES:  8
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  MARY A. APPOLLINA
CC        STREET:  P.O. BOX 2000, 126 E. LINCOLN AVENUE
CC        CITY:  RAHWAY
CC        STATE:  NJ
CC        COUNTRY:  USA
CC        ZIP:  07065
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US95/07295
CC        FILING DATE:
CC        CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  APPOLLINA, MARY A
CC        REGISTRATION NUMBER:  34,087
CC        REFERENCE/DOCKET NUMBER:  19244Y
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (908)594-3462
CC        TELEFAX:  (908)594-4720
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  2553 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  17..2221
SQ   Sequence 2553 BP; 749 A; 535 C; 610 G; 659 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 11;  Length 2553;
  Best Local Similarity  58.6%;  Pred. No. 1.01e+01;
  Matches    17;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    767 AAGTTGCTCTACAGGTTCCTGAAGTGGAA 795
          ||||| || |||        |||  ||||
Cp     33 aagttcctatacNNNNNNNgaataggaa 5


RESULT   18
ID   PCT-US93-05651-1 STANDARD; DNA; UNC; 6560 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9305651.
CC   Sequence 1, Application PC/TUS9305651
CC    GENERAL INFORMATION:
CC      TITLE OF INVENTION:  A Gene Which Prevents Programmed Cell Death
CC      NUMBER OF SEQUENCES:  5
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  diskette
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/05651
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  6560 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
```

```
CC       TOPOLOGY:  linear
CC     MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 6560 BP; 2040 A; 1270 C; 1207 G; 2023 T; 20 other;

  Query Match           50.0%;  Score 13;  DB 9;  Length 6560;
  Best Local Similarity  61.9%;  Pred. No. 1.01e+01;
  Matches    13;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    254 CTCCGCAAAGAATAGGAACTT 274
          |        ||||||||||||
Cp     22 cNNNNNNNNgaataggaactt 2


RESULT   19
ID   PCT-US94-13200-5 STANDARD; DNA; UNC; 7493 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application PC/TUS9413200.
CC   Sequence 5, Application PC/TUS9413200
CC    GENERAL INFORMATION:
CC      APPLICANT:  Emory University
CC      TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
CC      NUMBER OF SEQUENCES:  12
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Kilpatrick & Cody
CC        STREET:  1100 Peachtree Street, Suite 2800
CC        CITY:  Atlanta
CC        STATE:  Georgia
CC        COUNTRY:  US
CC        ZIP:  30309
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US94/13200
CC        FILING DATE:  15-NOV-1994
CC        CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Pabst, Patrea L.
CC        REGISTRATION NUMBER:  31,284
CC        REFERENCE/DOCKET NUMBER:  EMU106CIP(2)
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  404-815-6508
CC        TELEFAX:  404-815-6555
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7493 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA to mRNA
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Mus musculus
CC      FEATURE:
CC        NAME/KEY:  repeat_unit
CC        LOCATION:  1..407
CC        OTHER INFORMATION:  /rpt_type= "terminal"
CC        OTHER INFORMATION:  /note= "5'UTR"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
```



```
CC        LOCATION:  7471..7476
CC        OTHER INFORMATION:  /function= "PolyA_signal"
CC      FEATURE:
CC        NAME/KEY:  repeat_unit
CC        LOCATION:  7368..7493
CC        OTHER INFORMATION:  /rpt_type= "terminal"
CC        OTHER INFORMATION:  /note= "3'UTR"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  408..7367
CC        OTHER INFORMATION:  /product= "Coagulation Factor VIII"
CC      PUBLICATION INFORMATION:
CC        AUTHORS:  Elder, F.
CC        AUTHORS:  Lakich, D.
CC        AUTHORS:  Gitschier, J.
CC        TITLE:  Sequence of the Murine Factor VIII cDNA.
CC        JOURNAL:  Genomics
CC        VOLUME:  16
CC        PAGES:  374-379
CC        DATE:  1993
CC        RELEVANT RESIDUES IN SEQ ID NO:  5: FROM 1 TO 7476
SQ   Sequence 7493 BP; 2487 A; 1503 C; 1436 G; 2067 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 10;  Length 7493;
  Best Local Similarity  59.3%;  Pred. No. 1.01e+01;
  Matches    16;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   3714 GAAGATCCTATTCCACAAGATGAAGAG 3740
          |||| ||||||||        | | ||
Qy      1 gaagttcctattcNNNNNNNNgtatag 27


RESULT   20
ID   PCT-US94-13200-5 STANDARD; DNA; UNC; 7493 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application PC/TUS9413200.
CC   Sequence 5, Application PC/TUS9413200
CC    GENERAL INFORMATION:
CC      APPLICANT:  Emory University
CC      TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
CC      NUMBER OF SEQUENCES:  12
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Kilpatrick & Cody
CC        STREET:  1100 Peachtree Street, Suite 2800
CC        CITY:  Atlanta
CC        STATE:  Georgia
CC        COUNTRY:  US
CC        ZIP:  30309
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US94/13200
CC        FILING DATE:  15-NOV-1994
CC        CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Pabst, Patrea L.
CC        REGISTRATION NUMBER:  31,284
CC        REFERENCE/DOCKET NUMBER:  EMU106CIP(2)
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  404-815-6508
```

```
CC       TELEFAX:  404-815-6555
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7493 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA to mRNA
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Mus musculus
CC      FEATURE:
CC        NAME/KEY:  repeat_unit
CC        LOCATION:  1..407
CC        OTHER INFORMATION:  /rpt_type= "terminal"
CC        OTHER INFORMATION:  /note= "5'UTR"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  7471..7476
CC        OTHER INFORMATION:  /function= "PolyA_signal"
CC      FEATURE:
CC        NAME/KEY:  repeat_unit
CC        LOCATION:  7368..7493
CC        OTHER INFORMATION:  /rpt_type= "terminal"
CC        OTHER INFORMATION:  /note= "3'UTR"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  408..7367
CC        OTHER INFORMATION:  /product= "Coagulation Factor VIII"
CC      PUBLICATION INFORMATION:
CC        AUTHORS:  Elder, F.
CC        AUTHORS:  Lakich, D.
CC        AUTHORS:  Gitschier, J.
CC        TITLE:  Sequence of the Murine Factor VIII cDNA.
CC        JOURNAL:  Genomics
CC        VOLUME:  16
CC        PAGES:  374-379
CC        DATE:  1993
CC        RELEVANT RESIDUES IN SEQ ID NO:  5: FROM 1 TO 7476
SQ   Sequence 7493 BP; 2487 A; 1503 C; 1436 G; 2067 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 10;  Length 7493;
  Best Local Similarity  59.3%;  Pred. No. 1.01e+01;
  Matches    16;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   3714 GAAGATCCTATTCCACAAGATGAAGAG 3740
          |||| |||||| |        ||| ||
Cp     34 gaagttcctatacNNNNNNNNgaatag 8


RESULT   21
ID   PCT-US93-05705-1 STANDARD; DNA; UNC; 7653 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9305705.
CC   Sequence 1, Application PC/TUS9305705
CC    GENERAL INFORMATION:
CC      APPLICANT:  Massachusetts Institute of Technology
CC      TITLE OF INVENTION:  Inhibitors of Ced-3 and Related Proteins
CC      NUMBER OF SEQUENCES:  14
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Massachusetts Institute of Technology
CC        STREET:  77 Massachusetts Avenue
```



```
CC        CITY:  Cambridge
CC        STATE:  Massachusetts
CC        COUNTRY:  U.S.A.
CC        ZIP:  02139
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  diskette
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/05705
CC        FILING DATE:  19930714
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7653 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 7653 BP; 2429 A; 1455 C; 1271 G; 2498 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 9;  Length 7653;
  Best Local Similarity  60.0%;  Pred. No. 1.01e+01;
  Matches    15;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   4399 TATTCCATGAAGAGGATATGAACTT 4423
          |||||        | ||| ||||||
Qy      9 tattcNNNNNNNNgtataggaactt 33


RESULT   22
ID   PCT-US93-05701-18 STANDARD; DNA; UNC; 7653 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 18, Application PC/TUS9305701.
CC   Sequence 18, Application PC/TUS9305701
CC    GENERAL INFORMATION:
CC      APPLICANT:  Massachusetts Institute of Technology
CC      TITLE OF INVENTION:  Cloning and Characterization of Cell Death Genes
CC      NUMBER OF SEQUENCES:  29
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Massachusetts Institute of Technology
CC        STREET:  77 Massachusetts Avenue
CC        CITY:  Cambridge
CC        STATE:  Massachusetts
CC        COUNTRY:  U.S.A.
CC        ZIP:  02139
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  diskette
CC        COMPUTER:
CC        OPERATING SYSTEM:
CC        SOFTWARE:
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/05701
CC        FILING DATE:  19930614
CC        CLASSIFICATION:
CC    INFORMATION FOR SEQ ID NO:  18:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  7653 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 7653 BP; 2429 A; 1455 C; 1271 G; 2498 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 9;  Length 7653;
  Best Local Similarity  60.0%;  Pred. No. 1.01e+01;
```

```
  Matches    15;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   4399 TATTCCATGAAGAGGATATGAACTT 4423
          |||||        | ||| ||||||
Qy      9 tattcNNNNNNNNgtataggaactt 33


RESULT   23
ID   US-07-753-520B-4 STANDARD; DNA; UNC; 8316 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application US/07753520B.
CC   Sequence 4, Application US/07753520B
CC   Patent No. 5352595
CC    GENERAL INFORMATION:
CC      APPLICANT:  Tapscott,J.;Weintraub,H.M.;Palmer,T.D.
CC      TITLE OF INVENTION:  "MyoD REGULATORY REGION"
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Christensen, O'Connor, Johnson and Kindness
CC        STREET:  2800 Pacific First Center, 1420 Fifth Avenue
CC        CITY:  Seattle
CC        STATE:  Washington
CC        COUNTRY:  USA
CC        ZIP:  98101-2347
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Diskette-5.25 inch, 1.2Mb storage
CC        COMPUTER:  IBM PC/386 Compatible
CC        OPERATING SYSTEM:  MS-DOS 4.01
CC        SOFTWARE:  Word for Windows-t
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/753,520B
CC        FILING DATE:  19910903
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  none
CC        FILING DATE:  none
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Sundsmo,John,S.
CC        REGISTRATION NUMBER:  34,446
CC        REFERENCE/DOCKET NUMBER:  FHCR-1-5789
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  1-206-682-8100; 1-206-224-0727 (direct)
CC        TELEFAX:  1-206-224-0779
CC        TELEX:  4938023
CC     INFORMATION FOR SEQ ID NO:  4:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  8316 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  Other;plasmid DNA
CC        DESCRIPTION:  pLHDMDN-NSA: 5'LTR (position 1-1159); y+ (position
CC        DESCRIPTION:  1160-1640); HisD (position 1641-2928), Myo-D NSA ApaI f
  ragment
CC        DESCRIPTION:  (position 2929-4389); driving neo (position 4390-5259);
   with 3'LTR
CC        DESCRIPTION:  (position 5260-5964); Figures 8A-8C.
SQ   Sequence 8316 BP; 1922 A; 2246 C; 2255 G; 1880 T; 13 other;

  Query Match           50.0%;  Score 13;  DB 3;  Length 8316;
  Best Local Similarity  61.9%;  Pred. No. 1.01e+01;
  Matches    13;  Conservative    0;  Mismatches  8;  Indels   0;  Gaps   0;
```



```
Db    241 ATACATAACTGAGAATAGGAA 261
          ||||        |||||||||
Cp     25 atacNNNNNNNNgaataggaa 5


RESULT   24
ID   US-07-991-867B-1 STANDARD; DNA; UNC; 8457 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/07991867B.
CC   Sequence 1, Application US/07991867B
CC   Patent No. 5476781
CC    GENERAL INFORMATION:
CC      APPLICANT:  Moyer, Richard W.
CC      APPLICANT:  Hall, Richard L.
CC      APPLICANT:  Gruidl, Michael E.
CC      TITLE OF INVENTION:  No. 5476781el Entomopoxvirus Expression System
CC      NUMBER OF SEQUENCES:  66
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  David R. Saliwanchik
CC        STREET:  2421 N.W. 41st Street, Suite A-1
CC        CITY:  Gainesville
CC        STATE:  FL
CC        COUNTRY:  USA
CC        ZIP:  32606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/991,867B
CC        FILING DATE:  12-DEC-1992
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  WO 92/14818
CC        FILING DATE:  12-FEB-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/827,685
CC        FILING DATE:  30-JAN-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/657,584
CC        FILING DATE:  19-FEB-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Saliwanchik, David R.
CC        REGISTRATION NUMBER:  31,794
CC        REFERENCE/DOCKET NUMBER:  UF114.C3
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  904-375-8100
CC        TELEFAX:  904-372-5800
CC     INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  8457 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  DNA (genomic)
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Amsacta moorei entomopoxvirus
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  complement (65..1459)
CC      FEATURE:
CC        NAME/KEY:  CDS
```

```
CC        LOCATION:  1474..2151
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  complement (2239..2475)
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  2502..2987
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  3080..6091
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  complement (6277..6768)
SQ   Sequence 8457 BP; 3173 A; 951 C; 1006 G; 3327 T; 0 other;

  Query Match           50.0%;  Score 13;  DB 4;  Length 8457;
  Best Local Similarity  60.0%;  Pred. No. 1.01e+01;
  Matches    15;  Conservative    0;  Mismatches  10;  Indels    0;  Gaps    0;

Db   7080 AAGTTTCTATATATTTACACGAATA 7104
          ||||| |||||         |||||
Cp     33 aagttcctatacNNNNNNNNgaata 9


RESULT   25
ID   US-07-753-520B-3 STANDARD; DNA; UNC; 9115 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application US/07753520B.
CC   Sequence 3, Application US/07753520B
CC   Patent No. 5352595
CC    GENERAL INFORMATION:
CC      APPLICANT:  Tapscott,J.;Weintraub,H.M.;Palmer,T.D.
CC      TITLE OF INVENTION:  "MyoD REGULATORY REGION"
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Christensen, O'Connor, Johnson and Kindness
CC        STREET:  2800 Pacific First Center, 1420 Fifth Avenue
CC        CITY:  Seattle
CC        STATE:  Washington
CC        COUNTRY:  USA
CC        ZIP:  98101-2347
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Diskette-5.25 inch, 1.2Mb storage
CC        COMPUTER:  IBM PC/386 Compatible
CC        OPERATING SYSTEM:  MS-DOS 4.01
CC        SOFTWARE:  Word for Windows-t
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/753,520B
CC        FILING DATE:  19910903
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  none
CC        FILING DATE:  none
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Sundsmo,John,S.
CC        REGISTRATION NUMBER:  34,446
CC        REFERENCE/DOCKET NUMBER:  FHCR-1-5789
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  1-206-682-8100; 1-206-224-0727 (direct)
CC        TELEFAX:  1-206-224-0779
CC        TELEX:  4938023
CC     INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
```



```
CC        LENGTH:  9115 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  Other;plasmid DNA
CC        DESCRIPTION:  pLHDMDN-53: 5'LTR (position 1-1159); Y+ (position
CC        DESCRIPTION:  1159-1640); HisD (position 1641-3007); Myo-D 531.4 ApaI
   fragment
CC        DESCRIPTION:  (position 3008-5248); driving neo (position 5249-6117);
   with a
CC        DESCRIPTION:  3'LTR (position 6118-6823) coupled to a pBR322 plasmid
  (position
CC        DESCRIPTION:  6824-9115); Figures 7A-7D.
SQ   Sequence 9115 BP; 2183 A; 2408 C; 2474 G; 2036 T; 14 other;

  Query Match           50.0%;  Score 13;  DB 3;  Length 9115;
  Best Local Similarity  61.9%;  Pred. No. 1.01e+01;
  Matches    13;  Conservative    0;  Mismatches   8;  Indels    0;  Gaps    0;

Db    241 ATACATAACTGAGAATAGGAA 261
          ||||        |||||||||
Cp     25 atacNNNNNNNNgaataggaa 5


RESULT   26
ID   US-08-278-685-4 STANDARD; DNA; UNC; 31 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 4, Application US/08278685.
CC   Sequence 4, Application US/08278685
CC   Patent No. 5468483
CC    GENERAL INFORMATION:
CC      APPLICANT:  Thompson, Mark
CC      APPLICANT:  Gaertner, Frank H.
CC      TITLE OF INVENTION:  No. 5468483el Bacillus thuringiensis Isolate
CC      TITLE OF INVENTION:  Having Anti-Protozoan Activity
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Roman Saliwanchik
CC        STREET:  2421 N.W. 41st Street, Suite A-1
CC        CITY:  Gainesville
CC        STATE:  FL
CC        COUNTRY:  USA
CC        ZIP:  32606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/278,685
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/654,166
CC        FILING DATE:  12-FEB-1991
CC        APPLICATION NUMBER:  US 08/091,527
CC        FILING DATE:  12-AUG-1993
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Saliwanchik, Roman
CC        REGISTRATION NUMBER:  21,023
CC        REFERENCE/DOCKET NUMBER:  07/654,166
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  904-375-8100
```

```
CC        TELEFAX:  904-372-5800
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  31 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 31 BP; 6 A; 6 C; 10 G; 9 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 4;  Length 31;
  Best Local Similarity  59.1%;  Pred. No. 3.58e+01;
  Matches    13;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db      7 GAACTTCCTATTCGTGGTGGTG 28
          ||| |||||||||        |
Qy      1 gaagttcctattcNNNNNNNNg 22


RESULT   27
ID   PCT-US93-10443-9 STANDARD; DNA; UNC; 237 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 9, Application PC/TUS9310443.
CC   Sequence 9, Application PC/TUS9310443
CC    GENERAL INFORMATION:
CC      APPLICANT:  David D. Moore
CC      APPLICANT:  Jae W. Lee
CC      TITLE OF INVENTION:  NUCLEAR HORMONE RECEPTOR-
CC      TITLE OF INVENTION:  INTERACTING POLYPEPTIDES AND
CC      TITLE OF INVENTION:  RELATED MOLECULES AND METHODS
CC      NUMBER OF SEQUENCES:  30
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Fish & Richardson
CC        STREET:  225 Franklin Street
CC        CITY:  Boston
CC        STATE:  Massachusetts
CC        COUNTRY:  U.S.A.
CC        ZIP:  02110-2804
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
CC        COMPUTER:  IBM PS/2 Model 50Z or 55SX
CC        OPERATING SYSTEM:  MS-DOS (Version 5.0)
CC        SOFTWARE:  WordPerfect (Version 5.1)
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/10443
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  07/969,136
CC        FILING DATE:  October 30, 1992
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Paul T. Clark
CC        REGISTRATION NUMBER:  30,162
CC        REFERENCE/DOCKET NUMBER:  00786/099002
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (617) 542-5070
CC        TELEFAX:  (617) 542-8906
CC        TELEX:  200154
CC    INFORMATION FOR SEQ ID NO:  9:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  237
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
```



```
CC        TOPOLOGY:  linear
SQ   Sequence 237 BP; 83 A; 38 C; 52 G; 64 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 9;  Length 237;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db     77 TAAATGAAAGGGAACAGGAACTTC 100
          ||         ||| |||||||||
Cp     24 tacNNNNNNNNgaataggaacttc 1


RESULT   28
ID   US-08-026-320A-3 STANDARD; DNA; UNC; 360 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application US/08026320A.
CC   Sequence 3, Application US/08026320A
CC   Patent No. 5419904
CC    GENERAL INFORMATION:
CC      APPLICANT:  Irie, Reiko F
CC      TITLE OF INVENTION:  HUMAN B-LYMPHOBLASTOID CELL LINE
CC      TITLE OF INVENTION:  SECRETING ANTI-GANGLIOSIDE ANTIBODY
CC      NUMBER OF SEQUENCES:  11
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Poms, Smith, Lande & Rose
CC        STREET:  2029 Century Park East, Suite 3800
CC        CITY:  Los Angeles
CC        STATE:  California
CC        COUNTRY:  United States of America
CC        ZIP:  90067
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  WordPerfect 5.1
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/026,320A
CC        FILING DATE:  26-FEB-1993
CC        CLASSIFICATION:  424
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/609803
CC        FILING DATE:  05-NOV-1990
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Oldenkamp, David J
CC        REGISTRATION NUMBER:  29421
CC        REFERENCE/DOCKET NUMBER:  94268
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  3107885046
CC        TELEFAX:  3102771297
CC    INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  360 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Homo sapiens
CC        INDIVIDUAL ISOLATE:  Epstein Barr Virus transformed B
CC        INDIVIDUAL ISOLATE:  cell
CC        CELL TYPE:  B-cell
```

```
CC        CELL LINE:  L612
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  1..360
CC        OTHER INFORMATION:  /function= "Immunoglobulin light
CC        OTHER INFORMATION:  chain"
CC        OTHER INFORMATION:  /product= "HuMab L612 Light Chain Variable Region
  "
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  58..108
CC        OTHER INFORMATION:  /function= "Complementary
CC        OTHER INFORMATION:  determining region 1 (CDR1)"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  154..174
CC        OTHER INFORMATION:  /function= "Complementary
CC        OTHER INFORMATION:  determining region 2 (CDR2)"
CC      FEATURE:
CC        NAME/KEY:  misc_feature
CC        LOCATION:  271..297
CC        OTHER INFORMATION:  /function= "Complementary
CC        OTHER INFORMATION:  determining region 3 (CDR3)"
SQ   Sequence 360 BP; 88 A; 103 C; 86 G; 83 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 4;  Length 360;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db     77 TATACAGCTCCAACAATAAGAACT 100
          |||||         |||| |||||
Cp     26 tatacNNNNNNNNgaataggaact 3


RESULT   29
ID   PCT-US94-02629-54 STANDARD; DNA; UNC; 1229 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 54, Application PC/TUS9402629.
CC   Sequence 54, Application PC/TUS9402629
CC    GENERAL INFORMATION:
CC      APPLICANT:  King, Te-Piao
CC      TITLE OF INVENTION:  CLONING AND RECOMBINANT PRODUCTION OF
CC      TITLE OF INVENTION:  VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC      TITLE OF INVENTION:  HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED T
  HEREON
CC      NUMBER OF SEQUENCES:  62
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Klauber & Jackson
CC        STREET:  411 Hackensack Avenue
CC        CITY:  Hackensack
CC        STATE:  New Jersey
CC        COUNTRY:  USA
CC        ZIP:  07601
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US94/02629
CC        FILING DATE:  10-MAR-1994
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
```



```
CC        APPLICATION NUMBER:  US 08/180,209
CC        FILING DATE:  11-JAN-1994
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 08/031,400
CC        FILING DATE:  11-MAR-1993
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Jackson Esq., David A.
CC        REGISTRATION NUMBER:  26,742
CC        REFERENCE/DOCKET NUMBER:  600-1-074 PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  201 487-5800
CC        TELEFAX:  201 343-1684
CC        TELEX:  133521
CC    INFORMATION FOR SEQ ID NO:  54:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1229 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  61..1056
SQ   Sequence 1229 BP; 413 A; 229 C; 261 G; 326 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 10;  Length 1229;
  Best Local Similarity 100.0%;  Pred. No. 3.58e+01;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    349 AATAGGAACTTC 360
          ||||||||||||
Cp     12 aataggaacttc 1


RESULT   30
ID   US-08-133-347-3 STANDARD; DNA; UNC; 1635 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application US/08133347.
CC   Sequence 3, Application US/08133347
CC   Patent No. 5348888
CC    GENERAL INFORMATION:
CC      APPLICANT:  SHIRATORI,  Toshikazu
CC      APPLICANT:  INOUE,  Chihiro
CC      APPLICANT:  KITAGAWA,  Yoshichika
CC      APPLICANT:  kUSANO,  Tomonobu
CC      TITLE OF INVENTION:  DNA FRAGMENT CODING FOR MERCURIC REDUCTASE OF
CC      TITLE OF INVENTION:  THIOBACILLUS, AND RECOMBINANT PLASMID
CC      NUMBER OF SEQUENCES:  5
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Nixon & Vanderhye, P.C.
CC        STREET:  1100 No. 5348888th Glebe Road, 8th Floor
CC        CITY:  Arlington
CC        STATE:  Virginia
CC        COUNTRY:  U.S.A.
CC        ZIP:  22201
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy Disk
CC        COMPUTER:  IBM PC Compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
```

```
CC        APPLICATION NUMBER:  US/08/133,347
CC        FILING DATE:  08-OCT-1993
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  CRAWFORD, ARTHUR R
CC        REGISTRATION NUMBER:  25,327
CC        REFERENCE/DOCKET NUMBER:  159-30
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (703) 816-4000
CC        TELEFAX:  (703) 816-4100
CC    INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1635 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      ORIGINAL SOURCE:
CC        ORGANISM:  T. ferrooxidans strain E-15
CC      IMMEDIATE SOURCE:
CC        CLONE:  plasmid pTM314
CC      FEATURE:
CC        OTHER INFORMATION:  expresses T. ferrooxidans  merA
SQ   Sequence 1635 BP; 300 A; 517 C; 545 G; 273 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 3;  Length 1635;
  Best Local Similarity  57.1%;  Pred. No. 3.58e+01;
  Matches    16;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   1508 TTCGTAACCGGATGACGGTACAGGAACT 1535
           ||| ||  |        ||| |||||||
Qy      5 ttcctattcNNNNNNNNgtataggaact 32


RESULT   31
ID   US-08-277-540-2 STANDARD; DNA; UNC; 1749 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 2, Application US/08277540.
CC   Sequence 2, Application US/08277540
CC   Patent No. 5474901
CC    GENERAL INFORMATION:
CC      APPLICANT:  Drayna, Dennis T., Eaton, Dan L.
CC      TITLE OF INVENTION:  No. 5474901el Plasma Carboxypeptidase
CC      NUMBER OF SEQUENCES:  8
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Genentech, Inc.
CC        STREET:  460 Point San Bruno Blvd
CC        CITY:  South San Francisco
CC        STATE:  California
CC        COUNTRY:  USA
CC        ZIP:  94080
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  patin (Genentech)
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/277,540
CC        FILING DATE:  19-JUL-1994
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/167727
CC        FILING DATE:  15-DEC-1993
CC      PRIOR APPLICATION DATA:
```



```
CC        APPLICATION NUMBER:  07/959944
CC        FILING DATE:  14-OCT-1992
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  07/649591
CC        FILING DATE:  01-FEB-91
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Hasak, Janet E.
CC        REGISTRATION NUMBER:  28,616
CC        REFERENCE/DOCKET NUMBER:  689D1C1D1
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  415/225-1896
CC        TELEFAX:  415/952-9881
CC        TELEX:  910/371-7168
CC    INFORMATION FOR SEQ ID NO:  2:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  1749 bases
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
SQ   Sequence 1749 BP; 521 A; 361 C; 342 G; 525 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 4;  Length 1749;
  Best Local Similarity 100.0%;  Pred. No. 3.58e+01;
  Matches    12;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    805 GAATAGGAACTT 816
          ||||||||||||
Cp     13 gaataggaactt 2


RESULT   32
ID   US-08-217-327-5 STANDARD; DNA; UNC; 2230 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application US/08217327.
CC   Sequence 5, Application US/08217327
CC   Patent No. 5474925
CC    GENERAL INFORMATION:
CC      APPLICANT:  John, Maliyakal E
CC      APPLICANT:  Barton, Kenneth A
CC      TITLE OF INVENTION:  Immobilized Proteins in Cotton Fiber
CC      NUMBER OF SEQUENCES:  16
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Quarles and Brady
CC        STREET:  P.O. Box 2113
CC        CITY:  Madison
CC        STATE:  WI
CC        COUNTRY:  USA
CC        ZIP:  53701-2113
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/217,327
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/812,233
CC        FILING DATE:  19-DEC-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Seay, Nicholas J
CC        REGISTRATION NUMBER:  27,386
```

```
CC        REFERENCE/DOCKET NUMBER:  1122990831
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  608-251-5000
CC        TELEFAX:  608-251-9166
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  2230 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC        ORGANISM:  Daucus carota
CC      IMMEDIATE SOURCE:
CC        CLONE:  extensin gene
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  750..1670
CC      FEATURE:
CC        NAME/KEY:  sig_peptide
CC        LOCATION:  750..845
SQ   Sequence 2230 BP; 742 A; 541 C; 307 G; 640 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 4;  Length 2230;
  Best Local Similarity  59.1%;  Pred. No. 3.58e+01;
  Matches    13;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db   1854 AAGTTCATATACATTTCGAGGA 1875
            |||||| |||||        ||
Cp     33 aagttcctatacNNNNNNNNga 12


RESULT   33
ID   PCT-US95-07391A-1 STANDARD; DNA; UNC; 2339 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9507391A.
CC   Sequence 1, Application PC/TUS9507391A
CC    GENERAL INFORMATION:
CC      APPLICANT:  IBEX TECHNOLOGIES and
CC      APPLICANT:  ZIMMERMANN, Joseph
CC      TITLE OF INVENTION:  Nucleic Acid Sequences And Expression
CC      TITLE OF INVENTION:  Systems For Heparinase II And Heparinase III Deriv
  ed From
CC      TITLE OF INVENTION:  Flavobacterium heparinum
CC      NUMBER OF SEQUENCES:  26
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Hale and Dorr
CC        STREET:  1455 Pennsylvania Avenue, N.W.
CC        CITY:  Washington, D.C.
CC        COUNTRY:  U.S.A.
CC        ZIP:  20004
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US95/07391A
CC        FILING DATE:  09-JUNE-1995
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
```



```
CC        APPLICATION NUMBER:  08/258,639
CC        FILING DATE:  10 JUNE 1994
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  BAKER, Hollie L.
CC        REGISTRATION NUMBER:  31,321
CC        REFERENCE/DOCKET NUMBER:  104385.116PCT
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (202)942-8400
CC        TELEFAX:  (202)942-8484
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  2339 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 2339 BP; 667 A; 469 C; 571 G; 632 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 11;  Length 2339;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   2168 CTTTTCCGTTGTTGGTAAAGGAAC 2191
            || |||        ||| ||||||
Qy      8 ctattcNNNNNNNNgtataggaac 31


RESULT   34
ID   US-08-105-483-222 STANDARD; DNA; UNC; 2356 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 222, Application US/08105483.
CC   Sequence 222, Application US/08105483
CC   Patent No. 5494807
CC    GENERAL INFORMATION:
CC      APPLICANT:  Paoletti, Enzo
CC      TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
CC      TITLE OF INVENTION:  STRAIN
CC      NUMBER OF SEQUENCES:  462
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Curtis, Morris & Safford
CC        ADDRESSEE:  c/o William S. Frommer
CC        STREET:  530 Fifth Avenue
CC        CITY:  New York
CC        STATE:  NY
CC        COUNTRY:  USA
CC        ZIP:  10036
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/105,483
CC        FILING DATE:  12-AUG-1993
CC        CLASSIFICATION:  424
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/847,951
CC        FILING DATE:  06-MAR-1992
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Frommer, William S.
CC        REGISTRATION NUMBER:  25,506
CC        REFERENCE/DOCKET NUMBER:  454310-2400
CC      TELECOMMUNICATION INFORMATION:
```

```
CC        TELEPHONE:  (212) 840-3333
CC        TELEFAX:  (212) 840-0712
CC     INFORMATION FOR SEQ ID NO:  222:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  2356 base pairs
CC         TYPE:  nucleic acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
SQ   Sequence 2356 BP; 761 A; 397 C; 340 G; 858 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 4;  Length 2356;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    220 GAAATTCTTATACGTATCGGCGAA 243
          ||| ||| |||||        |||
Cp     34 gaagttcctatacNNNNNNNNgaa 11

RESULT   35
ID   PCT-US91-07035-11 STANDARD; DNA; UNC; 2679 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 11, Application PC/TUS9107035.
CC   Sequence 11, Application PC/TUS9107035
CC    GENERAL INFORMATION:
CC      APPLICANT:  Gelfand, David H.
CC      APPLICANT:  Abramson, Richard D.
CC      TITLE OF INVENTION:  5' TO 3' EXONUCLEASE MUTATIONS OF
CC      TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASES
CC      NUMBER OF SEQUENCES:  38
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Cetus Corporation
CC        STREET:  1400 Fifty-third Street
CC        CITY:  Emeryville
CC        STATE:  California
CC        ZIP:  94608
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  WordPerfect 5.0
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US91/07035
CC        FILING DATE:  19910930
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 590,490
CC        FILING DATE:  28-SEP-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 590,466
CC        FILING DATE:  28-SEP-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 590,213
CC        FILING DATE:  28-SEP-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 523,394
CC        FILING DATE:  15-MAY-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 143,441
CC        FILING DATE:  12-JAN-1988
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 063,509
CC        FILING DATE:  17-JUN-1987
CC      PRIOR APPLICATION DATA:
```



```
CC        APPLICATION NUMBER:  US 899,241
CC        FILING DATE:  22-AUG-1986
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 746,121
CC        FILING DATE:  15-AUG-1991
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  WO PCT/US90/07641
CC        FILING DATE:  21-DEC-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 585,471
CC        FILING DATE:  20-SEP-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 455,611
CC        FILING DATE:  22-DEC-1989
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 609,157
CC        FILING DATE:  02-NOV-1990
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 557,517
CC        FILING DATE:  24-JUL-1990
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Sias Ph.D, Stacey R.
CC        REGISTRATION NUMBER:  32,630
CC        REFERENCE/DOCKET NUMBER:  Case No. 2580
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  415-420-3300
CC     INFORMATION FOR SEQ ID NO:  11:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  2679 base pairs
CC         TYPE:  NUCLEIC ACID
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  DNA (genomic)
CC       HYPOTHETICAL:  NO
CC       ANTI-SENSE:  NO
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Thermosipho africanus
CC       FEATURE:
CC         NAME/KEY:  CDS
CC         LOCATION:  1..2676
SQ   Sequence 2679 BP; 1045 A; 295 C; 516 G; 823 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 7;  Length 2679;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   1869 TACTTTTCATCAAACAGGAACTTC 1892
           |||         || |||||||||
Cp     24 tacNNNNNNNNgaataggaacttc 1


RESULT   36
ID   PCT-US93-06251-79 STANDARD; DNA; UNC; 2757 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 79, Application PC/TUS9306251.
CC   Sequence 79, Application PC/TUS9306251
CC    GENERAL INFORMATION:
CC      APPLICANT:  Wickstrom, Eric and Rife, Jason P.
CC      TITLE OF INVENTION:  Trivalent Synthesis of Oligonucleotides Containing
CC      TITLE OF INVENTION:  Stereospecific Alkylphosphonates and Arylphosphona
  tes
CC      NUMBER OF SEQUENCES:  93
CC      CORRESPONDENCE ADDRESS:
```

```
CC        ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
CC        STREET:  400 Garden City Plaza
CC        CITY:  Garden City
CC        STATE:  NY
CC        COUNTRY:  USA
CC        ZIP:  11530
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/06251
CC        FILING DATE:  19930630
CC        CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  DiGiglio, Frank S.
CC        REGISTRATION NUMBER:  31,346
CC        REFERENCE/DOCKET NUMBER:  8586
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  516-742-4343
CC        TELEFAX:  516-742-4366
CC        TELEX:  230 901 SANS UR
CC    INFORMATION FOR SEQ ID NO:  79:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  2757 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 2757 BP; 693 A; 713 C; 809 G; 542 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 9;  Length 2757;
  Best Local Similarity  56.3%;  Pred. No. 3.58e+01;
  Matches    18;  Conservative    0;  Mismatches  14;  Indels   0;  Gaps   0;

Db    169 AGTTCCTGTGCCTTAAGAACATTAGAACCTTC 200
          ||||||| | |         | ||| | ||||
Cp     32 agttcctatacNNNNNNNNgaataggaacttc 1


RESULT   37
ID   PCT-US95-08071-111 STANDARD; DNA; UNC; 3033 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 111, Application PC/TUS9508071.
CC   Sequence 111, Application PC/TUS9508071
CC    GENERAL INFORMATION:
CC      APPLICANT:  Suzuki, Shintaro
CC      TITLE OF INVENTION:  Protocadherin Materials and Methods
CC      NUMBER OF SEQUENCES:  115
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray, &
CC        ADDRESSEE:  Borun
CC        STREET:  6300 Sears Tower, 233 S. Wacker Drive
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
```



```
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US95/08071
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/12588
CC        FILING DATE:  23 DEC 1993
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/998,003
CC        FILING DATE:  29 DEC 1992
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Noland, Greta E.
CC        REGISTRATION NUMBER:  35,302
CC        REFERENCE/DOCKET NUMBER:  32149
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  312/474-6300
CC        TELEFAX:  312/474-0448
CC        TELEX:  25-3856
CC    INFORMATION FOR SEQ ID NO:  111:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3033 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  138..2528
SQ   Sequence 3033 BP; 785 A; 723 C; 723 G; 802 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 11;  Length 3033;
  Best Local Similarity  57.7%;  Pred. No. 3.58e+01;
  Matches    15;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   2787 TCATATTCGCTGTAACGAATTGGAAC 2812
          || |||||        | || |||||
Qy      6 tcctattcNNNNNNNNgtataggaac 31


RESULT   38
ID   PCT-US95-08071-111 STANDARD; DNA; UNC; 3033 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 111, Application PC/TUS9508071.
CC   Sequence 111, Application PC/TUS9508071
CC    GENERAL INFORMATION:
CC      APPLICANT:  Suzuki, Shintaro
CC      TITLE OF INVENTION:  Protocadherin Materials and Methods
CC      NUMBER OF SEQUENCES:  115
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray, &
CC        ADDRESSEE:  Borun
CC        STREET:  6300 Sears Tower, 233 S. Wacker Drive
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US95/08071
```

```
CC        FILING DATE:
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/12588
CC        FILING DATE:  23 DEC 1993
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/998,003
CC        FILING DATE:  29 DEC 1992
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Noland, Greta E.
CC        REGISTRATION NUMBER:  35,302
CC        REFERENCE/DOCKET NUMBER:  32149
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  312/474-6300
CC        TELEFAX:  312/474-0448
CC        TELEX:  25-3856
CC    INFORMATION FOR SEQ ID NO:  111:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3033 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  138..2528
SQ   Sequence 3033 BP; 785 A; 723 C; 723 G; 802 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 11;  Length 3033;
  Best Local Similarity  57.7%;  Pred. No. 3.58e+01;
  Matches    15;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   2787 TCATATTCGCTGTAACGAATTGGAAC 2812
            || ||| |        |||| |||||
Cp     29 tcctatacNNNNNNNNgaataggaac 4


RESULT   39
ID   PCT-US92-11337-5 STANDARD; DNA; UNC; 3513 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 5, Application PC/TUS9211337.
CC   Sequence 5, Application PC/TUS9211337
CC    GENERAL INFORMATION:
CC      APPLICANT:  PAYNE, JEWEL M.
CC      APPLICANT:  HICKLE, LESLIE A.
CC      TITLE OF INVENTION:  NOVEL BACILLUS THURINGIENSIS ISOLATES
CC      TITLE OF INVENTION:  ACTIVE AGAINST PHTHIRAPTERA PESTS
CC      NUMBER OF SEQUENCES:  16
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  DAVID R. SALIWANCHIK
CC        STREET:  2421 N.W. 41st STREET, SUITE A-1
CC        CITY:  GAINESVILLE
CC        STATE:  FL
CC        COUNTRY:  USA
CC        ZIP:  32606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/11337
CC        FILING DATE:  19921231
```



```
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  97/828,788
CC        FILING DATE:
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  SALIWANCHIK, DAVID R.
CC        REGISTRATION NUMBER:  31,794
CC        REFERENCE/DOCKET NUMBER:  MA75
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  904-375-8100
CC        TELEFAX:  904-372-5800
CC    INFORMATION FOR SEQ ID NO:  5:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3513 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC        ORGANISM:  BACILLUS THURINGIENSIS
CC        STRAIN:  KENYAE
CC        INDIVIDUAL ISOLATE:  PS81F
CC      IMMEDIATE SOURCE:
CC        LIBRARY:  LAMBDAGEM (TM) - 11 LIBRARY OF AUGUST SICK
CC        CLONE:  81F
SQ   Sequence 3513 BP; 1169 A; 592 C; 769 G; 983 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 8;  Length 3513;
  Best Local Similarity  59.1%;  Pred. No. 3.58e+01;
  Matches    13;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db   1744 GAACTTCCTATTCGTGGTGGTG 1765
            ||| |||||||||        |
Qy      1 gaagttcctattcNNNNNNNNg 22


RESULT   40
ID   US-08-278-685-1 STANDARD; DNA; UNC; 3513 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application US/08278685.
CC   Sequence 1, Application US/08278685
CC   Patent No. 5468483
CC    GENERAL INFORMATION:
CC      APPLICANT:  Thompson, Mark
CC      APPLICANT:  Gaertner, Frank H.
CC      TITLE OF INVENTION:  No. 5468483el Bacillus thuringiensis Isolate
CC      TITLE OF INVENTION:  Having Anti-Protozoan Activity
CC      NUMBER OF SEQUENCES:  4
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Roman Saliwanchik
CC        STREET:  2421 N.W. 41st Street, Suite A-1
CC        CITY:  Gainesville
CC        STATE:  FL
CC        COUNTRY:  USA
CC        ZIP:  32606
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
```

```
CC        APPLICATION NUMBER:  US/08/278,685
CC        FILING DATE:
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/654,166
CC        FILING DATE:  12-FEB-1991
CC        APPLICATION NUMBER:  US 08/091,527
CC        FILING DATE:  12-AUG-1993
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Saliwanchik, Roman
CC        REGISTRATION NUMBER:  21,023
CC        REFERENCE/DOCKET NUMBER:  07/654,166
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  904-375-8100
CC        TELEFAX:  904-372-5800
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  3513 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
SQ   Sequence 3513 BP; 1169 A; 592 C; 769 G; 983 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 4;  Length 3513;
  Best Local Similarity  59.1%;  Pred. No. 3.58e+01;
  Matches    13;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db   1744 GAACTTCCTATTCGTGGTGGTG 1765
            ||| |||||||||        |
Qy      1 gaagttcctattcNNNNNNNNg 22


RESULT   41
ID   PCT-US92-03222-38 STANDARD; DNA; UNC; 4131 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 38, Application PC/TUS9203222.
CC   Sequence 38, Application PC/TUS9203222
CC    GENERAL INFORMATION:
CC      APPLICANT:  Beavo, Joseph A.
CC      APPLICANT:  Bentley, Kelley
CC      APPLICANT:  Charbonneau, Harry
CC      APPLICANT:  Sonnenburg, William K.
CC      TITLE OF INVENTION:  DNA Encoding Mammalian
CC      TITLE OF INVENTION:  Phosphodiesterases
CC      NUMBER OF SEQUENCES:  58
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
CC        ADDRESSEE:  Bicknell
CC        STREET:  Two First National Plaza, 20 South Clark
CC        STREET:  Street
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60603
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US92/03222
CC        FILING DATE:  19920420
```



```
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/688,356
CC        FILING DATE:  04-APR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Noland, Greta E.
CC        REGISTRATION NUMBER:  35,302
CC        REFERENCE/DOCKET NUMBER:  27866/30822
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (312) 346-5750
CC        TELEFAX:  (312) 984-9740
CC        TELEX:  25-3856
CC    INFORMATION FOR SEQ ID NO:  38:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  4131 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  148..2910
SQ   Sequence 4131 BP; 866 A; 1233 C; 1174 G; 858 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 8;  Length 4131;
  Best Local Similarity  60.0%;  Pred. No. 3.58e+01;
  Matches    12;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   1742 TCCTATACAAGAAAGTGAAT 1761
            ||||||||        ||||
Cp     29 tcctatacNNNNNNNNgaat 10


RESULT   42
ID   US-07-872-644-38 STANDARD; DNA; UNC; 4131 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 38, Application US/07872644.
CC   Sequence 38, Application US/07872644
CC   Patent No. 5389527
CC    GENERAL INFORMATION:
CC      APPLICANT:  Beavo, Joseph A.
CC      APPLICANT:  Bentley, Kelley
CC      APPLICANT:  Charbonneau, Harry
CC      APPLICANT:  Sonnenburg, William K.
CC      TITLE OF INVENTION:  DNA Encoding Mammalian
CC      TITLE OF INVENTION:  Phosphodiesterases
CC      NUMBER OF SEQUENCES:  58
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
CC        ADDRESSEE:  Bicknell
CC        STREET:  Two First National Plaza, 20 South Clark
CC        STREET:  Street
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60603
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/07/872,644
```

```
CC        FILING DATE:  19920420
CC        CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  US 07/688,356
CC        FILING DATE:  04-APR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  No. 5389527and, Greta E.
CC        REGISTRATION NUMBER:  35,302
CC        REFERENCE/DOCKET NUMBER:  27866/30822
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (312) 346-5750
CC        TELEFAX:  (312) 984-9740
CC        TELEX:  25-3856
CC    INFORMATION FOR SEQ ID NO:  38:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  4131 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  148..2910
SQ   Sequence 4131 BP; 866 A; 1233 C; 1174 G; 858 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 3;  Length 4131;
  Best Local Similarity  60.0%;  Pred. No. 3.58e+01;
  Matches    12;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   1742 TCCTATACAAGAAAGTGAAT 1761
            ||||||||        ||||
Cp     29 tcctatacNNNNNNNNgaat 10


RESULT   43
ID   PCT-US95-07744A-15 STANDARD; DNA; UNC; 4146 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 15, Application PC/TUS9507744A.
CC   Sequence 15, Application PC/TUS9507744A
CC    GENERAL INFORMATION:
CC      APPLICANT:  Trustees of The University of Pennsylvania
CC      TITLE OF INVENTION:  Plant Genes for Sensitivity to Ethylene
CC      TITLE OF INVENTION:  and Pathogens
CC      NUMBER OF SEQUENCES:  82
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz & Norris
CC        STREET:  One Liberty Place, 46th floor
CC        CITY:  Philadelphia
CC        STATE:  PA
CC        COUNTRY:  USA
CC        ZIP:  19103
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US95/07744A
CC        FILING DATE:  15-JUNE-1995
CC        CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER:  08/261,822
CC        FILING DATE:  June 17, 1994
```

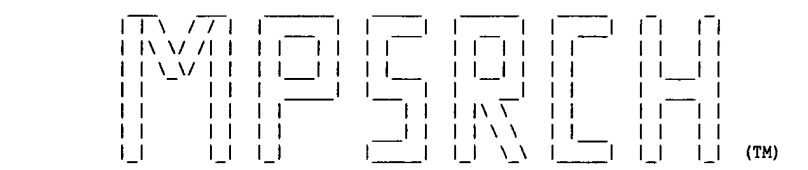


```
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Beardell, Lori Y.
CC        REGISTRATION NUMBER:  34,293
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (215) 568-3100
CC        TELEFAX:  (215) 568-3439
CC    INFORMATION FOR SEQ ID NO:  15:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  4146 base pairs
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
SQ   Sequence 4146 BP; 1265 A; 707 C; 744 G; 1430 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 11;  Length 4146;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   1610 TATACAAAAAATAAAATGGGAACT 1633
            |||||          ||| ||||||
Cp     26 tatacNNNNNNNNgaataggaact 3


RESULT   44
ID   US-08-045-806-3 STANDARD; DNA; UNC; 5261 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 3, Application US/08045806.
CC   Sequence 3, Application US/08045806
CC   Patent No. 5378822
CC    GENERAL INFORMATION:
CC      APPLICANT:  Bradfield, Christopher Alan
CC      APPLICANT:  Dolwick, Kristin Marie
CC      APPLICANT:  Poland, Alan
CC      TITLE OF INVENTION:  Ah Receptor cDNA and Method of
CC      TITLE OF INVENTION:  Determining Human Risks To Environmental Pollutant
  s
CC      NUMBER OF SEQUENCES:  23
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Tilton, Fallon, Lungmus & Chestnut
CC        STREET:  100 South Wacker Drive, Suite 960
CC        CITY:  Chicago
CC        STATE:  Illinois
CC        COUNTRY:  USA
CC        ZIP:  60606-4002
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  US/08/045,806
CC        FILING DATE:  19930408
CC        CLASSIFICATION:  435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Fentress, Susan B.
CC        REGISTRATION NUMBER:  31,327
CC        REFERENCE/DOCKET NUMBER:  NU-9207
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  (312)-456-8000
CC        TELEFAX:  (312)-456-7776
```

```
CC    INFORMATION FOR SEQ ID NO:  3:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  5261 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  cDNA
CC      FEATURE:
CC        NAME/KEY:  CDS
CC        LOCATION:  383..2927
SQ   Sequence 5261 BP; 1625 A; 1102 C; 976 G; 1558 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 3;  Length 5261;
  Best Local Similarity  58.3%;  Pred. No. 3.58e+01;
  Matches    14;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    842 ATACATCAGAGTGTATATGAACTT 865
          ||||        | ||| ||||||
Cp     25 atacNNNNNNNNgaataggaactt 2


RESULT   45
ID   PCT-US93-03076-1 STANDARD; DNA; UNC; 8298 BP.
AC   xxxxxx
DT   01-JAN-1900
DE   Sequence 1, Application PC/TUS9303076.
CC   Sequence 1, Application PC/TUS9303076
CC    GENERAL INFORMATION:
CC      APPLICANT:  Whitehead Institute for Biomedical Research
CC      TITLE OF INVENTION:  GAP-Associated Protein p190 and
CC      TITLE OF INVENTION:  Transduction
CC      NUMBER OF SEQUENCES:  20
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
CC        STREET:  2 Militia Drive
CC        CITY:  Lexington
CC        STATE:  MA
CC        COUNTRY:  US
CC        ZIP:  02173
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE:  Floppy disk
CC        COMPUTER:  IBM PC compatible
CC        OPERATING SYSTEM:  PC-DOS/MS-DOS
CC        SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER:  PCT/US93/03076
CC        FILING DATE:  19930331
CC        CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME:  Granahan, Patricia
CC        REGISTRATION NUMBER:  32,227
CC        REFERENCE/DOCKET NUMBER:  WHI92-03A
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE:  617-861-6240
CC        TELEFAX:  617-861-9540
CC    INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  8298 base pairs
CC        TYPE:  NUCLEIC ACID
CC        STRANDEDNESS:  double
CC        TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      FEATURE:
CC        NAME/KEY:  CDS
```



```
CC        LOCATION:  731..5272
SQ   Sequence 8298 BP; 2180 A; 2086 C; 2039 G; 1993 T; 0 other;

  Query Match           46.2%;  Score 12;  DB 9;  Length 8298;
  Best Local Similarity  59.1%;  Pred. No. 3.58e+01;
  Matches    13;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db   1231 CAGGGGCATGAACAGGAACTTC 1252
          |        ||| |||||||||
Cp     22 cNNNNNNNNgaataggaacttc 1


Search completed: Tue May 14 11:40:38 1996
Job time : 10 secs.
```



Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue May 14 13:49:31 1996; MasPar time 32.48 Seconds
741.130 Million cell updates/sec

Tabular output not generated.

Title: >FLP
Description: (1-34) from frt.seq
Perfect Score: 26
N.A. Sequence: 1 gaagttcctattcNNNNNNNNgtataggaacttc 34
Comp: cttcaaggataagNNNNNNNNcatatccttgaag

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 264399 seqs, 353985056 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-new11
1:BCT 2:FUN 3:INV1 4:INV2 5:INV3 6:MAM 7:ORG 8:PLN
9:PRI1 10:PRI2 11:PRI3 12:PRO1 13:PRO2 14:ROD 15:SYN
16:UNC 17:VRT 18:VIR

Database: genbank91
19:BCT1 20:BCT2 21:BCT3 22:BCT4 23:BCT5 24:BCT6 25:BCT7
26:INV1 27:INV2 28:INV3 29:INV4 30:INV5 31:MAM1 32:MAM2
33:PAT1 34:PAT2 35:PAT3 36:PHG 37:PLN1 38:PLN2 39:PLN3
40:PLN4 41:PLN5 42:PLN6 43:PLN7 44:PRI1 45:PRI2 46:PRI3
47:PRI4 48:PRI5 49:PRI6 50:PRI7 51:PRI8 52:PRI9 53:ROD1
54:ROD2 55:ROD3 56:ROD4 57:ROD5 58:ROD6 59:ROD7 60:STR
61:SYN 62:UNA 63:VRL1 64:VRL2 65:VRL3 66:VRL4 67:VRL5
68:VRL6 69:VRT1 70:VRT2 71:VRT3

Database: genbank-new11
72:BCT1 73:BCT2 74:INV1 75:INV2 76:MAM 77:PHG 78:PLN
79:PRI1 80:PRI2 81:PRI3 82:ROD 83:STR 84:SYN 85:UNA
86:VRL 87:VRT

Database: u-embl44_91
88:part1

Statistics: Mean 7.033; Variance 2.584; scale 2.722

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	26	100.0	68	35	I08526	Sequence 1 from paten	4.89e-08
	2	26	100.0	121	61	APDNATSR2	Artificial plasmid DN	4.89e-08
	3	26	100.0	125	61	APDNATSR3	Artificial plasmid DN	4.89e-08
	4	26	100.0	154	61	APDNATSR1	Artificial plasmid DN	4.89e-08
	5	26	100.0	200	43	YSCPL2M	Yeast (S.cerevisiae)	4.89e-08
c	6	26	100.0	1019	41	SCPLA1	Part of the 2 micron	4.89e-08
c	7	26	100.0	1578	41	SCOR01	Yeast sequence contai	4.89e-08
	8	26	100.0	1814	61	CVPMAKK76	Cloning vector pMAKK7	4.89e-08
	9	26	100.0	1888	61	CVPMAKC76	Cloning vector pMAKC7	4.89e-08
c	10	26	100.0	5616	61	PRS424	Yeast episomal vector	4.89e-08
c	11	26	100.0	5726	61	PRS426	Yeast episomal vector	4.89e-08
c	12	26	100.0	5797	61	PRS423	Yeast episomal vector	4.89e-08
c	13	26	100.0	5807	61	CVPFL45L	multicopy Saccharomyc	4.89e-08
c	14	26	100.0	6010	43	YSCTRAM1	Cloning vector pYSVE2	4.89e-08
c	15	26	100.0	6037	43	YSCTRAM2	Cloning vector pYADE4	4.89e-08
c	16	26	100.0	6063	61	CVPFL44L	multicopy Saccharomyc	4.89e-08
	17	26	100.0	6318	41	SCA21	2 micron plasmid of y	4.89e-08
	18	26	100.0	6318	43	YSCPLASM	Yeast (S.cerevisiae)	4.89e-08
c	19	26	100.0	6318	43	YSCPLASM	Yeast (S.cerevisiae)	4.89e-08
c	20	26	100.0	6318	41	SCA21	2 micron plasmid of y	4.89e-08
c	21	26	100.0	6445	61	SYNECOYST	Cloning vector sequen	4.89e-08
	22	26	100.0	6624	2	CV37458	Yeast CUP1 expression	4.89e-08
	23	26	100.0	6624	84	CVU37458	Yeast CUP1 expression	4.89e-08
c	24	26	100.0	6849	61	PRS425	Yeast episomal vector	4.89e-08
	25	26	100.0	7769	61	SYNYEP24V	YEp24 yeast extrachro	4.89e-08
c	26	26	100.0	7822	61	CVPFL46L	multicopy Saccharomyc	4.89e-08
c	27	26	100.0	7834	15	CV33753	Yeast episomal clonin	4.89e-08
c	28	26	100.0	7834	61	CVU33753	Yeast episomal clonin	4.89e-08
	29	26	100.0	7859	34	A17115	Yeast expression vect	4.89e-08
	30	26	100.0	7859	35	I13185	Sequence 4 from paten	4.89e-08
	31	26	100.0	7859	88	A19996	SEQ ID NO: 4; Synthet	4.89e-08
	32	26	100.0	7984	34	A18079	yeast expression vect	4.89e-08
	33	26	100.0	8117	61	CVU29899	Cloning vector pACT2	4.89e-08
c	34	26	100.0	8393	15	CV30497	Cloning vector pAS2-1	4.89e-08
c	35	26	100.0	8393	61	CVU30497	Cloning vector pAS2-1	4.89e-08
c	36	26	100.0	10667	61	YEP213	Yeast episomal vector	4.89e-08
c	37	26	100.0	10667	61	YEP13	Yeast episomal vector	4.89e-08
c	38	22	84.6	200	43	YSCPL2M	Yeast (S.cerevisiae)	1.34e-04
	39	22	84.6	5616	61	PRS424	Yeast episomal vector	1.34e-04
	40	22	84.6	5726	61	PRS426	Yeast episomal vector	1.34e-04
c	41	22	84.6	6624	84	CVU37458	Yeast CUP1 expression	1.34e-04
c	42	22	84.6	6624	2	CV37458	Yeast CUP1 expression	1.34e-04
	43	22	84.6	7834	15	CV33753	Yeast episomal clonin	1.34e-04
c	44	22	84.6	7859	88	A19996	SEQ ID NO: 4; Synthet	1.34e-04
c	45	22	84.6	7984	34	A18079	yeast expression vect	1.34e-04

ALIGNMENTS

RESULT 1
LOCUS I08526 68 bp DNA PAT 14-NOV-1994
DEFINITION Sequence 1 from patent WO 8703006.
ACCESSION I08526
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 68)

```
  AUTHORS   Rogers,D.T. and Szostak,J.W.
  TITLE     YEAST STRAINS
  JOURNAL   Patent: WO 8703006-A 1 21-MAY-1987;

COMMENT     NCBI gi: 588764
FEATURES             Location/Qualifiers
     source          1..68
                     /organism="unknown"
BASE COUNT       18 a     15 c     13 g     22 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 35;  Length 68;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     25 gaagttcctattctctagaaagtataggaacttc 58
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    2
LOCUS       APDNATSR2     121 bp    DNA             SYN       21-JUN-1995
DEFINITION  Artificial plasmid DNA containing target site for specific
            recombinase (121 bp).
ACCESSION   X87981
KEYWORDS    beta-galactosidase; recombinase target site.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 121)
  AUTHORS   Snaith,M.R., Kilby,N.J. and Murray,A.H.
  TITLE     An E. coli system for assay of FLP site-specific recombination on
            substrate plasmids
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 121)
  AUTHORS   Snaith,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUN-1995) to the EMBL/GenBank/DDBJ databases. M.
            Snaith, University of Cambridge, Dept of Genetics, Downing Site,
            Downing Street, Cambridge CB2 3Eh, UK
COMMENT     NCBI gi: 870842
FEATURES             Location/Qualifiers
     source          1..121
                     /organism="Artificial sequences"
                     /note="plasmid DNA"
     misc_binding    23..70
                     /note="FRT target site"
                     /bound_moiety="FLP site-specific recombinase"
                     /evidence=experimental
     misc_feature    71..121
                     /note="modified portion fo beta-galactosidase"
BASE COUNT       27 a     31 c     29 g     34 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 121;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     37 gaagttcctattctctagaaagtataggaacttc 70
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    3
```



```
LOCUS       APDNATSR3     125 bp    DNA             SYN       21-JUN-1995
DEFINITION  Artificial plasmid DNA containing target site for specific
            recombinase (125 bp).
ACCESSION   X87982
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 125)
  AUTHORS   Snaith,M.R., Kilby,N.J. and Murray,A.H.
  TITLE     An E. coli system for assay of FLP site-specific recombination on
            substrate plasmids
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 125)
  AUTHORS   Snaith,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUN-1995) to the EMBL/GenBank/DDBJ databases. M.
            Snaith, University of Cambridge, Dept of Genetics, Downing Site,
            Downing Street, Cambridge CB2 3Eh, UK
COMMENT     NCBI gi: 870843
FEATURES             Location/Qualifiers
     source          1..125
                     /organism="Artificial sequences"
     CDS             56..>125
                     /note="pid:e;  NCBI gi: 870844"
                     /codon_start=1
                     /product="beta-galactosidase"
                     /translation="MEKLLFRSSYSLESIGTSSLALA"
     misc_feature    56..125
                     /note="5' modified portion fo beta-galactosidase"
     misc_binding    61..108
                     /note="FRT target site"
                     /bound_moiety="FLP site-specific recombinase"
                     /evidence=experimental
BASE COUNT       30 a     27 c     33 g     35 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 125;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     75 gaagttcctattctctagaaagtataggaacttc 108
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    4
LOCUS       APDNATSR1     154 bp    DNA             SYN       21-JUN-1995
DEFINITION  Artificial plasmid DNA containing target site for specific
            recombinase (154 bp).
ACCESSION   X87980
KEYWORDS    beta-galactosidase; recombinase target site.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 154)
  AUTHORS   Snaith,M.R., Kilby,N.J. and Murray,A.H.
  TITLE     An E. coli system for assay of FLP site-specific recombination on
            substrate plasmids
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 154)
  AUTHORS   Snaith,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUN-1995) to the EMBL/GenBank/DDBJ databases. M.
```

```
            Snaith, University of Cambridge, Dept of Genetics, Downing Site,
            Downing Street, Cambridge CB2 3Eh, UK
COMMENT     NCBI gi: 870840
FEATURES             Location/Qualifiers
     source          1..154
                     /organism="Artificial sequences"
                     /note="plasmid DNA"
     CDS             56..>118
                     /note="pid:e;  NCBI gi: 870841"
                     /codon_start=1
                     /product="beta-galactosidase"
                     /translation="MEKLLFRSSYSLESIGTSERF"
     misc_feature    56..154
                     /note="5' modified portion fo beta-galactosidase"
     misc_binding    61..108
                     /note="FRT target site"
                     /bound_moiety="FLP site-specific recombinase"
                     /evidence=experimental
BASE COUNT       41 a     37 c     37 g     39 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 154;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     75 gaagttcctattctctagaaagtataggaacttc 108
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    5
LOCUS       YSCPL2M       200 bp    DNA             PLN       10-DEC-1984
DEFINITION  Yeast (S.cerevisiae) 2 micron plasmid (A-form) inverted repeat
            region.
ACCESSION   K01710
KEYWORDS    plasmid.
SOURCE      Yeast (Saccharomyces cerevisiae) 2 micron plasmid DNA.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Endomycetales;
            Saccharomycetaceae.
REFERENCE   1  (bases 1 to 200)
  AUTHORS   Fagrelius,T.J. and Livingston,D.M.
  TITLE     Location of DNAase I sensitive cleavage sites in the yeast 2 mu-m
            plasmid DNA chromosome
  JOURNAL   J. Mol. Biol. 173, 1-13 (1984)
  MEDLINE   84138647
COMMENT     [1] examines whether cleavage sites are specific when the
            DNA-associated protein is stripped away and draws the conclusion
            that the specificity of DNAase I is dependent on the presence of
            nucleoprotein.

            NCBI gi: 172188
FEATURES             Location/Qualifiers
     source          1..200
                     /organism="Saccharomyces cerevisiae"
BASE COUNT       57 a     47 c     46 g     50 t
ORIGIN      103 bp upstream of XbaI site.

  Query Match          100.0%;  Score 26;  DB 43;  Length 200;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     90 gaagttcctattctctagaaagtataggaacttc 123
          ||||||||||||||        ||||||||||||||
```



```
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    6
LOCUS       SCPLA1       1019 bp    DNA             PLN       06-JUL-1989
DEFINITION  Part of the 2 micron plasmid of yeast encompassing one of the
            inverted repeats.
ACCESSION   V01322
KEYWORDS    terminal repeat.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Eumycota; Ascomycotina; Hemiascomycetes; Saccharomycetales;
            Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   Hindley,J. and Phear,G.A.
  TITLE     Sequence of 1019 nucleotides encompassing one of the inverted
            repeats from the yeast 2 micrometer plasmid
  JOURNAL   Nucleic Acids Res. 7 (2), 361-375 (1979)
  MEDLINE   80034481
COMMENT     KST SCE.PLASMID (INCOMPL.).

            NCBI gi: 4181
FEATURES             Location/Qualifiers
     source          1..1019
                     /organism="Saccharomyces cerevisiae"
                     /plasmid="2 micron plasmid"
BASE COUNT      271 a    192 c    225 g    330 t      1 others
ORIGIN

  Query Match          100.0%;  Score 26;  DB 41;  Length 1019;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    299 gaagttcctatactttctagagaataggaacttc 332
          ||||||||||||        ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT    7
LOCUS       SCOR01       1578 bp    DNA             PLN       13-JUL-1983
DEFINITION  Yeast sequence containing a replication origin.
ACCESSION   V01317
KEYWORDS    origin of replication.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Eumycota; Ascomycotina; Hemiascomycetes; Saccharomycetales;
            Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 1578)
  AUTHORS   Hindley,J. and Phear,G.A.
  TITLE     Sequencing long DNA fragments cloned in bacteriophage M13 by using
            internal primers. The sequence analysis of a yeast DNA fragment
            containing a replication origin
  JOURNAL   Biochem. J. 199 (3), 819-823 (1981)
  MEDLINE   82182087
COMMENT     NCBI gi: 4083
FEATURES             Location/Qualifiers
     source          1..1578
                     /organism="Saccharomyces cerevisiae"
BASE COUNT      445 a    301 c    306 g    526 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 41;  Length 1578;
```

```
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   1199 gaagttcctatactttctagagaataggaacttc 1232
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1

RESULT    8
LOCUS       CVPMAKK76    1814 bp    DNA   circular  SYN       16-SEP-1994
DEFINITION  Cloning vector pMAKK76 with kanamycin phosphotransferase (KnR)
            gene, complete sequence.
ACCESSION   U08460
KEYWORDS    .
SOURCE      Cloning vector pMAKK76.
  ORGANISM  Cloning vector pMAKK76
            Artificial sequence; Cloning vector.
REFERENCE   1  (bases 1 to 1814)
  AUTHORS   Posfai,G., Koob,M., Hradecna,Z., Hasan,N., Filutowicz,M. and
            Szybalski,W.
  TITLE     In vivo excision and amplification of large segments of the
            Escherichia coli genome
  JOURNAL   Nucleic Acids Res. 22 (12), 2392-2398 (1994)
  MEDLINE   94310070
REFERENCE   2  (bases 1 to 1814)
  AUTHORS   Posfai,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-1994) Gyorgy Posfai, University of Wisconsin,
            McArdle Laboratory for Cancer Research, 1400 University Avenue,
            Madison, WI 53706, USA
COMMENT     NCBI gi: 475708
FEATURES             Location/Qualifiers
     source          1..1814
                     /organism="Cloning vector pMAKK76"
                     /lab_host="Escherichia coli"
                     /plasmid=""
     misc_feature    17..50
                     /note="FRT site from yeast 2 micron plasmid"
     misc_feature    103
                     /note="T7 RNA polymerase transcription initiation site"
     misc_feature    112..168
                     /standard_name="multiple cloning site"
     misc_feature    complement(174)
                     /note="SP6 RNA polymerase transcription initiation site"
     rep_origin      430..805
                     /note="gamma replication origin from R6K"
                     /direction=LEFT
     CDS             complement(869..1663)
                     /gene="KnR"
                     /note="NCBI gi: 475709"
                     /codon_start=1
                     /function="kanamycin resistance"
                     /evidence=experimental
                     /transl_table=11
                     /product="kanamycin phosphotransferase"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
BASE COUNT      463 a    459 c    444 g    448 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 1814;
```



```
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     17 gaagttcctattctctagaaagtataggaacttc 50
          ||||||||||||        ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34

RESULT    9
LOCUS       CVPMAKC76    1888 bp    DNA   circular  SYN       16-SEP-1994
DEFINITION  Cloning vector pMAKC76 with chloramphenicol acetyltransferase (CmR)
            gene, complete sequence.
ACCESSION   U08461
KEYWORDS    .
SOURCE      Cloning vector pMAKC76.
  ORGANISM  Cloning vector pMAKC76
            Artificial sequence; Cloning vector.
REFERENCE   1  (bases 1 to 1814)
  AUTHORS   Posfai,G., Koob,M., Hradecna,Z., Hasan,N., Filutowicz,M. and
            Szybalski,W.
  TITLE     In vivo excision and amplification of large segments of the
            Escherichia coli genome
  JOURNAL   Nucleic Acids Res. 22 (12), 2392-2398 (1994)
  MEDLINE   94310070
REFERENCE   2  (bases 1 to 1888)
  AUTHORS   Posfai,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-1994) Gyorgy Posfai, University of Wisconsin,
            McArdle Laboratory for Cancer Research, 1400 University Avenue,
            Madison, WI 53706, USA
COMMENT     NCBI gi: 475710
FEATURES             Location/Qualifiers
     source          1..1888
                     /organism="Cloning vector pMAKC76"
                     /lab_host="Escherichia coli"
                     /plasmid=""
     misc_feature    17..50
                     /note="FRT site from yeast 2 micron plasmid"
     misc_feature    103
                     /note="T7 RNA polymerase transcription initiation site"
     misc_feature    112..168
                     /standard_name="multiple cloning site"
     misc_feature    complement(174)
                     /note="SP6 RNA polymerase transcription initiation site"
     rep_origin      430..805
                     /note="gamma replication origin from R6K"
                     /direction=LEFT
     CDS             complement(930..1589)
                     /gene="CmR"
                     /note="NCBI gi: 475711"
                     /codon_start=1
                     /function="chloramphenicol resistance"
                     /evidence=experimental
                     /transl_table=11
                     /product="chloramphenicol acetyltransferase"
                     /translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAF
                     LKTVKKNKHKFYPAFIHILARLMNAHPKFRMAMKDGELVIWDSVHPCYTVFHEQTETF
                     SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNV
                     ANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGG
                     A"
BASE COUNT      527 a    399 c    409 g    553 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 1888;
```

```
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     17 gaagttcctattctctagaaagtataggaacttc 50
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT  10
LOCUS       PRS424       5616 bp    DNA   circular  SYN       24-MAY-1995
DEFINITION  Yeast episomal vector pRS424 with TRP1 marker, complete sequence.
ACCESSION   U03453
KEYWORDS    .
SOURCE      Cloning vector pRS424.
  ORGANISM  Cloning vector pRS424
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 5616)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 5616)
  AUTHORS   Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and
            Hieter,P.
  TITLE     Multifunctional yeast high-copy-number shuttle vectors
  JOURNAL   Gene 110 (1), 119-122 (1992)
  MEDLINE   92184105
REFERENCE   3  (bases 1 to 5616)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416324
FEATURES             Location/Qualifiers
     source          1..5616
                     /organism="cloning vector pRS424"
BASE COUNT     1513 a   1221 c   1356 g   1526 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 5616;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5169 gaagttcctatactttctagagaataggaacttc 5202
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT  11
LOCUS       PRS426       5726 bp    DNA   circular  SYN       24-MAY-1995
DEFINITION  Yeast episomal vector pRS426 with URA3 marker, complete sequence.
ACCESSION   U03451
KEYWORDS    .
SOURCE      Cloning vector pRS426.
  ORGANISM  Cloning vector pRS426
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 5726)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
```



```
REFERENCE   2  (bases 1 to 5726)
  AUTHORS   Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and
            Hieter,P.
  TITLE     Multifunctional yeast high-copy-number shuttle vectors
  JOURNAL   Gene 110 (1), 119-122 (1992)
  MEDLINE   92184105
REFERENCE   3  (bases 1 to 5726)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416322
FEATURES             Location/Qualifiers
     source          1..5726
                     /organism="cloning vector pRS426"
BASE COUNT     1568 a   1246 c   1370 g   1542 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 5726;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5279 gaagttcctatactttctagagaataggaacttc 5312
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT  12
LOCUS       PRS423       5797 bp    DNA   circular  SYN       24-MAY-1995
DEFINITION  Yeast episomal vector pRS423 with HIS3 marker, complete sequence.
ACCESSION   U03454
KEYWORDS    .
SOURCE      Cloning vector pRS423.
  ORGANISM  Cloning vector pRS423
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 5797)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 5797)
  AUTHORS   Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and
            Hieter,P.
  TITLE     Multifunctional yeast high-copy-number shuttle vectors
  JOURNAL   Gene 110 (1), 119-122 (1992)
  MEDLINE   92184105
REFERENCE   3  (bases 1 to 5797)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416325
FEATURES             Location/Qualifiers
     source          1..5797
                     /organism="cloning vector pRS423"
BASE COUNT     1536 a   1308 c   1374 g   1579 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 5797;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;
```

```
Db   5351 gaagttcctatactttctagagaataggaacttc 5384
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1
```

```
RESULT   13
LOCUS       CVPFL45L     5807 bp    DNA             SYN       15-AUG-1995
DEFINITION  multicopy Saccharomyces cerevisiae/E. coli shuttle vector.
ACCESSION   X70267
KEYWORDS    2-micron yeast replication origin; pUC19 plasmid;
            TRP1 selectable marker.
SOURCE      cloning vectors.
  ORGANISM  cloning vectors
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 644; 1484 to 1610; 3862 to 5807)
  AUTHORS   Yanisch-Perron,C., Vieira,J. and Messing,J.
  TITLE     Improved M13 phage cloning vectors and host strains: nucleotide
            sequences of the M13mp18 and pUC19 vectors
  JOURNAL   Gene 33 (1), 103-119 (1985)
  MEDLINE   85180545
REFERENCE   2  (bases 644 to 1484)
  AUTHORS   Struhl,K., Stinchcomb,D.T., Scherer,S. and Davis,R.W.
  TITLE     High-frequency transformation of yeast: autonomous replication of
            hybrid DNA molecules
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 76 (3), 1035-1039 (1979)
  MEDLINE   79180126
REFERENCE   3  (bases 1610 to 3862)
  AUTHORS   Chevalier,M.R. and Lacroute,F.
  TITLE     Trancriptional and traductional expression of a chimeric
            bacterial- yeast plasmid in yeast
  JOURNAL   Gene 1, 11-19 (1980)
REFERENCE   4  (bases 1 to 5807)
  AUTHORS   Ozier-Kalogeropoullos,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-1993) to the EMBL/GenBank/DDBJ databases.
            Ozier-Kalogeropoullos O., CGM, CNRS, 91190 Gif sur Yvette, France
            e-mail:odile%FRCGM51.BITNET@vm.gmd.de
REFERENCE   5  (bases 1 to 5807)
  AUTHORS   Bonneaud,N., Ozier-Kalogeropoulos,O., Li,G.Y., Labouesse,M.,
            Minvielle-Sebastia,L. and Lacroute,F.
  TITLE     A family of low and high copy replicative, integrative and
            single-stranded S. cerevisiae/E. coli shuttle vectors
  JOURNAL   Yeast 7 (6), 609-615 (1991)
  MEDLINE   92116645
COMMENT     The pFL45L was constructed from pUC19 plasmid where two alu I sites
            were modified.The site 629 was replaced by a BglII linker and the
            site 747 by a ClaI site. The yeast selectable marker has been
            cloned in the BglII site and the 2 micron 2.2 kb EcoRI fragment
            containing ORI and STB gene has been cloned at the ClaI site. The
            pFL45L is described in Bonneaud et al (1991): A family of low and
            high copy replicative, integrative and single-stranded
            S.cerevisiae/E.coli shuttle vectors. YEAST, 7, 609-615.

            NCBI gi: 397132
FEATURES              Location/Qualifiers
     source           1..5807
                      /organism="Cloning vector"
BASE COUNT     1555 a   1261 c   1344 g   1647 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 5807;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;
```



```
Db   3139 gaagttcctatactttctagagaataggaacttc 3172
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1
```

```
RESULT   14
LOCUS       YSCTRAM1     6010 bp    DNA             PLN       18-JUL-1994
DEFINITION  Cloning vector pYSVE2 TRP1 and AMPr genes, complete cds.
ACCESSION   M74015
KEYWORDS    AmpR gene; TRP1 gene.
SOURCE      pYSVE2.
  ORGANISM  Synthetic construct
            Synthetic construct; Artificial sequences.
REFERENCE   1  (bases 1 to 6010)
  AUTHORS   Brunelli,J.P. and Pall,M.L.
  TITLE     A series of yeast vectors for expression of cDNAs and other DNA
            sequences
  JOURNAL   Yeast 9, 1299-1308 (1993)
  MEDLINE   94205259
REFERENCE   2  (bases 1 to 6010)
  AUTHORS   Pall,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1991) Martin L. Pall, Department of Genetics and
            Cell Biology, Washington State University, Pullman, WA 99164-4234,
            USA
COMMENT     NCBI gi: 173015
FEATURES              Location/Qualifiers
     source           1..6010
                      /organism="pYSVE2"
                      /note="cloning vector"
     CDS              complement(1868..2728)
                      /gene="AMPr"
                      /note="putative;  NCBI gi: 173016"
                      /codon_start=1
                      /transl_table=11
                      /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                      IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                      YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                      DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                      LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                      EIGASLIKHW"
     CDS              3041..3715
                      /gene="TRP1"
                      /note="putative;  NCBI gi: 173017"
                      /codon_start=1
                      /transl_table=11
                      /translation="MSVINFTGSSGPLVKVCGLQSTEAAECALDSDADLLGIICVPNR
                      KRTIDPVIARKISSLVKAYKNSSGTPKYLVGVFRNQPKEDVLALVNDYGIDIVQLHGD
                      ESWQEYQEFLGLPVIKRLVFPKDCNILLSAASQKPHSFIPLFDSEAGGTGELLDWNSI
                      SDWVGRQESPESLHFMLAGGLTPENVGDALRLNGVIGVDVSGGVETNGVKDSNKIANF
                      VKNAKK"
BASE COUNT     1629 a   1300 c   1385 g   1696 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 43;  Length 6010;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;
```

```
Db   5039 gaagttcctatactttctagagaataggaacttc 5072
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1
```

```
RESULT   15
LOCUS       YSCTRAM2      6037 bp    DNA             PLN       18-JUL-1994
DEFINITION  Cloning vector pYADE4 TRP1 and AMPr genes, complete cds.
ACCESSION   M74016
KEYWORDS    AMPr gene; TRP1 gene.
SOURCE      pYADE4.
  ORGANISM  Synthetic construct
            Synthetic construct; Artificial sequences.
REFERENCE   1  (bases 1 to 6037)
  AUTHORS   Brunelli,J.P. and Pall,M.L.
  TITLE     A series of yeast shuttle vectors for expression of cDNAs and other
            DNA sequences
  JOURNAL   Yeast 9 (12), 1299-1308 (1993)
  MEDLINE   94205259
REFERENCE   2  (bases 1 to 6037)
  AUTHORS   Pall,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1991) Martin L. Pall, Department of Genetics and
            Cell Biology, Washington State University, Pullman, WA 99164-4234,
            USA
COMMENT     NCBI gi: 173018
FEATURES             Location/Qualifiers
     source          1..6037
                     /organism="pYADE4"
                     /note="cloning vector"
     CDS             complement(1895..2755)
                     /gene="AMPr"
                     /note="putative;  NCBI gi: 387897"
                     /codon_start=1
                     /transl_table=11
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     CDS             3041..3742
                     /gene="TRP1"
                     /note="putative. start;  NCBI gi: 912422"
                     /codon_start=1
                     /transl_except=(pos:3041..3043,aa:OTHER)
                     /transl_table=11
                     /translation="MKHTKAAWSMSVINFTGSSGPLVKVCGLQSTEAAECALDSDADL
                     LGIICVPNRKRTIDPVIARKISSLVKAYKNSSGTPKYLVGVFRNQPKEDVLALVNDYG
                     IDIVQLHGDESWQEYQEFLGLPVIKRLVFPKDCNILLSAASQKPHSFIPLFDSEAGGT
                     GELLDWNSISDWVGRQESPESLHFMLAGGLTPENVGDALRLNGVIGVDVSGGVETNGV
                     KDSNKIANFVKNAKK"
BASE COUNT     1667 a   1270 c   1359 g   1741 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 43;  Length 6037;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5066 gaagttcctatactttctagagaataggaacttc 5099
          |||||||||||||        |||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   16
LOCUS       CVPFL44L      6063 bp    DNA             SYN       15-AUG-1995
DEFINITION  multicopy Saccharomyces cerevisiae/E. coli shuttle vector.
ACCESSION   X70484
KEYWORDS    2-micron yeast replication origin; pUC19 plasmid;
```



```
            URA3 selectable marker.
SOURCE      cloning vectors.
  ORGANISM  cloning vectors
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 641; 1743 to 1863; 4121 to 6063)
  AUTHORS   Yanisch-Perron,C., Vieira,J. and Messing,J.
  TITLE     Improved M13 phage cloning vectors and host strains: nucleotide
            sequences of the M13mp18 and pUC19 vectors
  JOURNAL   Gene 33 (1), 103-119 (1985)
  MEDLINE   85180545
REFERENCE   2  (bases 644 to 1740)
  AUTHORS   Bach,M.L., Lacroute,F. and Botstein,D.
  TITLE     Evidence for transcriptional regulation of orotidine-5'-phosphate
            decarboxylase in yeast by hybridization of mRNA to the yeast
            structural gene cloned in Escherichia coli
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 76 (1), 386-390 (1979)
  MEDLINE   79137106
REFERENCE   3  (bases 1866 to 4118)
  AUTHORS   Chevalier,M.R. and Lacroute,F.
  TITLE     Trancriptional and traductional expression of a chimeric
            bacterial- yeast plasmid in yeast
  JOURNAL   Gene 1, 11-19 (1980)
REFERENCE   4  (bases 1 to 6063)
  AUTHORS   Ozier-Kalogeropoullos,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-1993) to the EMBL/GenBank/DDBJ databases.
            Ozier-Kalogeropoullos O., CGM, CNRS, 91190 Gif sur Yvette, France
            e-mail:odile%FRCGM51.BITNET@vm.gmd.de
REFERENCE   5  (bases 1 to 6063)
  AUTHORS   Bonneaud,N., Ozier-Kalogeropoulos,O., Li,G.Y., Labouesse,M.,
            Minvielle-Sebastia,L. and Lacroute,F.
  TITLE     A family of low and high copy replicative, integrative and
            single-stranded S. cerevisiae/E. coli shuttle vectors
  JOURNAL   Yeast 7 (6), 609-615 (1991)
  MEDLINE   92116645
COMMENT     The pFL44L was constructed from pUC19 plasmid where two alu I sites
            were modified.The site 629 was replaced by a BglII linker and the
            site 747 by a ClaI site. The yeast selectable marker has been
            cloned in the BglII site and the 2 micron 2.2 kb EcoRI fragment
            containing ORI and STB gene has been cloned at the ClaI site. The
            pFL44L is described in Bonneaud et al (1991): A family of low and
            high copy replicative, integrative and single-stranded
            S.cerevisiae/E.coli shuttle vectors. YEAST, 7, 609-615.

            NCBI gi: 312626
FEATURES             Location/Qualifiers
     source          1..6063
                     /organism="Cloning vector"
BASE COUNT     1607 a   1385 c   1307 g   1764 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 6063;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3395 gaagttcctatactttctagagaataggaacttc 3428
          |||||||||||||        |||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   17
LOCUS       SCA21         6318 bp    DNA   circular  PLN       29-JUN-1995
DEFINITION  2 micron plasmid of yeast (circularly closed).
ACCESSION   V01323 J01347 L00321 L00322 L00323 L00324 M10185 M11111 M11593
```

```
            M14239 M14240 M14241 M14242 M14243 M14244 M14245 M14253 M14254
            M14255 M14256 M14257 M14258 M14259 M14591 M14592 M14593 M14594
            M14596 M14597 M14598
KEYWORDS    circular; origin of replication.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryotae; mitochondrial eukaryotes; Eumycota; Ascomycotina;
            Hemiascomycetes; Saccharomycetales; Saccharomycetaceae;
            Saccharomyces.
REFERENCE   1  (bases 1 to 6318)
  AUTHORS   Hartley,J.L. and Donelson,J.E.
  TITLE     Nucleotide sequence of the yeast plasmid
  JOURNAL   Nature 2869 , 860-865 (1980)
  MEDLINE   81012161
COMMENT     NCBI gi: 4182
FEATURES             Location/Qualifiers
     source          1..6318
                     /organism="Saccharomyces cerevisiae"
                     /plasmid="2 micron plasmid"
     CDS             complement(887..2008)
                     /note="NCBI gi: 4183"
                     /codon_start=1
                     /product="protein Baker"
                     /translation="MNGERLLACIKQCIMQHFQPMVYDESRCVIETTRGTFPVPDNYK
                     KYKTLAFAFVCGHVLNTDDTPVIEKELDWPDPALVYNTIVDRIINHPELSQFISVAFIS
                     QLKATIGEGLDINVKGTLNRRGKGIRRPKGVFFRYMESPFVNTKVTAFFSYLRDYNKI
                     ASEYHNNTKFILTFSCQAYWASGPNFSALKNVIRCSIIHEYISKFVEREQDKGHIGDQ
                     ELPPEEDPSRELNNVQHEVNSLTEQDAEADEGLWGEIDSLCEKWQSEAEDQTEAEIIA
                     DRIIGNSQRMANLKIRRTKFKSVLYHILKELIQSQGTVKVYRGSSFSHDSIKISLHYE
                     EQHITAVWVYLTVKFEEHWKPVDVEVEFRCKFKERKVDG"
     CDS             complement(4308..5198)
                     /note="NCBI gi: 4184"
                     /codon_start=1
                     /product="protein Charlie"
                     /translation="MDDIETAKNLTVKARTAYSVWDVCRLFIEMIAPDVDIDIESKRK
                     SDELLFPGYVIRPMESLTTGRPYGLDSSAEDSSVSSDSSAEVILPAAKMVKERFDSIG
                     NGMLSSQEASQAAIDLMLQNNKLLDNRKQLYKSIAIIIGRLPEKDKKRATEMLMRKMD
                     CTQLLVPPAPTEEDVMKLVSVVTQLLTLVPPDRQAALIGDLFIPESLKDIFNSFNELA
                     AENRLQQKKSELEGRTEVNHANTNEEVPSRRTRSRDTNARGAYKLQNTITEGPKAVPT
                     KKRRVATRVRGRKSRNTSRV"
BASE COUNT     1876 a    1284 c    1179 g    1979 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 41;  Length 6318;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    690 gaagttcctattctctagaaagtataggaacttc 723
          |||||||||||||        |||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT  18
LOCUS       YSCPLASM     6318 bp    DNA   circular  PLN       31-JUL-1992
DEFINITION  Yeast (S.cerevisiae) 2 micron circle plasmid, complete genome.
ACCESSION   J01347 L00321 L00322 L00323 L00324 M10185 M11111 M11593 M14239
            M14240 M14241 M14242 M14243 M14244 M14245 M14253 M14254 M14255
            M14256 M14257 M14258 M14259 M14591 M14592 M14593 M14594 M14595
            M14596 M14597 M14598 V01323
KEYWORDS    DNA-binding protein; Rep-1 protein; Rep-2 protein; circular;
            complete genome; d protein; plasmid; protein FLP; recombinase;
            repeat region.
SOURCE      Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71, p82-6B,
            CV20, pMMD2, pGP20, pJFS166 (see comment).
```



```
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Endomycetales;
            Saccharomycetaceae.
REFERENCE   1  (bases 1 to 1022)
  AUTHORS   Hindley,J. and Phear,G.A.
  TITLE     Sequence of 1019 nucleotides encompassing one of the inverted
            repeats from the yeast 2 micron plasmid
  JOURNAL   Nucleic Acids Res. 7, 361-375 (1979)
  MEDLINE   80034481
REFERENCE   2  (bases 1 to 6318; 1 to 6318)
  AUTHORS   Hartley,J.L. and Donelson,J.E.
  TITLE     Nucleotide sequence of the yeast plasmid
  JOURNAL   Nature 286, 860-865 (1980)
  MEDLINE   81012161
REFERENCE   3  (bases 3891 to 3990)
  AUTHORS   Broach,J.R., Guarascio,V.R. and Jayaram,M.
  TITLE     Recombination within the yeast plasmid 2-micron circle is
            site-specific
  JOURNAL   Cell 29, 227-234 (1982)
  MEDLINE   82259368
REFERENCE   4  (bases 3881 to 4020)
  AUTHORS   McLeod,M., Volkert,F. and Broach,J.R.
  TITLE     Components of the site-specific recombination system encoded by the
            yeast plasmid 2-micron circle
  JOURNAL   Cold Spring Harb. Symp. Quant. Biol. 49, 779-787 (1984)
  MEDLINE   85153059
REFERENCE   5  (bases 670 to 732)
  AUTHORS   Andrews,B.J., Proteau,G.A., Beatty,L.G. and Sadowski,P.D.
  TITLE     The FLP recombinase of the 2 micron circle DNA of yeast:
            Interaction with its target sequences
  JOURNAL   Cell 40, 795-803 (1985)
  MEDLINE   85176933
REFERENCE   6  (bases 5570 to 5605)
  AUTHORS   Babineau,D., Vetter,D., Andrews,B.J., Gronostajski,R.M.,
            Proteau,G.A., Beatty,L.G. and Sadowski,P.D.
  TITLE     The FLP protein of the 2-micron plasmid of yeast: Purification of
            the protein from Escherichia coli cells expressing the cloned FLP
            gene
  JOURNAL   J. Biol. Chem. 260, 12313-12319 (1985)
  MEDLINE   86008307
REFERENCE   7  (sites)
  AUTHORS   Gronostajski,R.M. and Sadowski,P.D.
  TITLE     Determination of DNA sequences essential for FLP-mediated
            recombination by a novel method
  JOURNAL   J. Biol. Chem. 260, 12320-12327 (1985)
  MEDLINE   86008308
REFERENCE   8  (sites)
  AUTHORS   Sutton,A. and Broach,J.R.
  TITLE     Signals for transcription initiation and termination in the
            Saccharomyces cerevisiae plasmid 2 micron circle
  JOURNAL   Mol. Cell. Biol. 5, 2770-2780 (1985)
  MEDLINE   86284639
REFERENCE   9  (sites)
  AUTHORS   Gronostajski,R.M. and Sadowski,P.D.
  TITLE     The FLP recombinase of the Saccharomyces cerevisiae 2-micron
            plasmid attaches covalently to DNA via a phosphotyrosyl linkage
  JOURNAL   Mol. Cell. Biol. 5, 3274-3279 (1985)
  MEDLINE   86310798
REFERENCE   10 (bases 667 to 739)
  AUTHORS   Senecoff,J.F., Bruckner,R.C. and Cox,M.M.
  TITLE     The FLP recombinase of the yeast 2-micron-m plasmid:
            Characterization of its recombination site
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82, 7270-7274 (1985)
  MEDLINE   86042647
```

```
REFERENCE   11 (sites)
  AUTHORS   McLeod,M., Craft,S. and Broach,J.R.
  TITLE     Identification of the crossover site during FLP-mediated
            recombination in the Saccharomyces cerevisiae plasmid 2 micron
            circle
  JOURNAL   Mol. Cell. Biol. 6, 3357-3367 (1986)
  MEDLINE   87089667
COMMENT     {8} sites; mRNA CAP sites and poly-adenylation sites. {9} sites;
            FLP binding.
            {7} sites; FLP cleavage.
            {11} sites; FLP-mediated recombination crossover site. Draft entry
            and clean copy sequence for [5] kindly provided by J.Senecoff,
            24-JAN-1986.
            Yeast 2 micron plasmid contains two 599 bp inverted repeats
            separated by a large unique (UL) and a small unique (US) region.
            During recombination the UL and US regions invert producing two
            sequence forms that differ in the orientation of one unique region
            relative to the other.  The A form is presented below. FLP is the
            only 2-micron circle-encoded protein needed for specific site
            recombination between the IRs of 2-micron circle.  The minimal size
            of the recombination site required for efficient FLP
            recombinase-catalyzed recombination in vitro is no more than 28 bp,
            which includes parts of two 13 bp inverted repeats (positions
            690-702 and 711-723) and all of an 8 bp spacer (703-710) [5].  The
            FLP recombinase cleaves the DNA at the boundaries of the spacer and
            becomes covalently linked to the spacer DNA [5],[9].  The
            efficiency of the recombination is reduced if the spacer in a
            recombinant site is increased or decreased by 1 bp, while the
            spacer in the second site is unaltered [5].  Recombination between
            two sites with identical 1-base pair additions or deletions is
            relatively unaffected, suggesting that pairing of sequences in the
            spacer regions is important in FLP-promoted recombination events
            [5].  The sequence asymmetry utilized by the recombinase to
            determine the orientation of the site is located uniquely within
            the spacer region. Another 13 bp direct repeat, is found at
            positions 676-688 [5].  FLP-mediated recombination involving two
            FLP sites that are inverted with respect to each other results in
            inversion of the DNA sequences between the sites [4].  If the
            participating recombination sites are in direct orientation, FLP
            promotes only the excision of the intervening DNA sequences [4].
            The Rep 1 and Rep proteins are involved plasmid partitioning and
            protein stability.
            A start codon in phase with the Rep1 coding region is located at
            positions 1966-1964.  Two CAP sites for Rep1 mRNA are located
            beyond the 'atg' codon (position 2008) at positions 2004 and 2005.
            Complete source information:
            Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71 [1],
            p82-6B [2], CV20 [3], pMMD2 [4], pGP20 [5], pJFS166 [10].

            NCBI gi: 172190
FEATURES            Location/Qualifiers
     source         1..6318
                    /organism="Saccharomyces cerevisiae"
     exon           1..545
     conflict       replace((157.160)..(157.160),"")
                    /citation=[1]
     conflict       replace((289.290)..(289.290),"")
                    /citation=[1]
     repeat_region  341..939
                    /note="IR2"
     conflict       replace((464.466)..(464.466),"")
                    /citation=[1]
     conflict       replace(558,"")
                    /citation=[1]
```



```
     conflict       replace(561,"")
                    /citation=[1]
     conflict       replace((622.624)..(622.624),"")
                    /citation=[1]
     conflict       replace(642,"")
                    /citation=[1]
     conflict       replace((665.666)..(665.666),"")
                    /citation=[1]
     misc_binding   673..722
                    /note="FLP recombinase binding site A [9]"
                    /bound_moiety="FLP recombinase"
     conflict       replace((793.794)..(793.794),"")
                    /citation=[1]
     mRNA           complement(836..2038)
                    /note="Repl mRNA (alt.)"
     mRNA           complement(836..2017)
                    /note="Repl mRNA (alt.)"
     mRNA           complement(836..2019)
                    /note="Repl mRNA (alt.)"
     mRNA           complement(836..2010)
                    /note="Repl mRNA (alt.)"
     mRNA           complement(836..2004)
                    /note="Repl mRNA (alt.)"
     mRNA           complement(836..2005)
                    /note="Repl mRNA (alt.)"
     CDS            complement(887..2008)
                    /note="Rep 1 protein;  NCBI gi: 172192"
                    /codon_start=1
                    /translation="MNGERLLACIKQCIMQHFQPMVYDESRCVIETTRGTFPVPDNYK
                    KYKTLAFAFVGHVLNTDDTPVIEKELDWPDPALVYNTIVDRIINHPELSQFISVAFIS
                    QLKATIGEGLDINVKGTLNRRGKGIRRPKGVFFRYMESPFVNTKVTAFFSYLRDYNKI
                    ASEYHNNTKFILTFSCQAYWASGPNFSALKNVIRCSIIHEYISKFVEREQDKGHIGDQ
                    ELPPEEDPSRELNNVQHEVNSLTEQDAEADEGLWGEIDSLCEKWQSEAEDQTEAEIIA
                    DRIIGNSQRMANLKIRRTKFKSVLYHILKELIQSQGTVKVYRGSSFSHDSIKISLHYE
                    EQHITAVWVYLTVKFEEHWKPVDVEVEFRCKFKERKVDG"
     mRNA           2254..2841
                    /note="D mRNA (alt.; 5' end +/- 3 bp)"
     mRNA           2254..2861
                    /note="D mRNA (alt.; 5' end +/- 3 bp)"
     CDS            2271..2816
                    /note="D protein;  NCBI gi: 172193"
                    /codon_start=1
                    /translation="MPYKTAIDCIEELATQCFLSKLTDDDVSTFRRVCSKENDIIKLA
                    LRIPRTIDYTSILRLLYDTLPLRSLSFNEALPLFCYSIDPAQQRQCDLRFYLRDVVKL
                    ARPRKRLEMQKALLQWLPSLLSDVTLQLLNDIRIRFEEIQPNIRQTVLQIYDRTCYPS
                    LNFEHPNLGVFPETDSIFEPV"
     repeat_region  3714..4312
                    /note="IR1"
     misc_binding   3930..3979
                    /note="FLP-recombinase binding site B [9]"
                    /bound_moiety="FLP recombinase"
     mRNA           complement(4108..5182)
                    /note="REP2 mRNA (major alt.)"
     mRNA           complement(4108..5183)
                    /note="REP2 mRNA (major alt.)"
     mRNA           complement(4108..5184)
                    /note="REP2 mRNA (major alt.)"
     mRNA           complement(4108..5223)
                    /note="REP2 mRNA (minor alt.)"
     mRNA           complement(4108..5195)
                    /note="REP2 mRNA (major alt.)"
...
Note: remainder of annotations omitted.
```

```
  Query Match          100.0%;  Score 26;  DB 43;  Length 6318;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    690 gaagttcctattctctagaaagtataggaacttc 723
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   19
LOCUS       YSCPLASM     6318 bp    DNA   circular  PLN       31-JUL-1992
DEFINITION  Yeast (S.cerevisiae) 2 micron circle plasmid, complete genome.
ACCESSION   J01347 L00321 L00322 L00323 L00324 M10185 M11111 M11593 M14239
            M14240 M14241 M14242 M14243 M14244 M14245 M14253 M14254 M14255
            M14256 M14257 M14258 M14259 M14591 M14592 M14593 M14594 M14595
            M14596 M14597 M14598 V01323
KEYWORDS    DNA-binding protein; Rep-1 protein; Rep-2 protein; circular;
            complete genome; d protein; plasmid; protein FLP; recombinase;
            repeat region.
SOURCE      Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71, p82-6B,
            CV20, pMMD2, pGP20, pJFS166 (see comment).
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Endomycetales;
            Saccharomycetaceae.
REFERENCE   1  (bases 1 to 1022)
  AUTHORS   Hindley,J. and Phear,G.A.
  TITLE     Sequence of 1019 nucleotides encompassing one of the inverted
            repeats from the yeast 2 micron plasmid
  JOURNAL   Nucleic Acids Res. 7, 361-375 (1979)
  MEDLINE   80034481
REFERENCE   2  (bases 1 to 6318; 1 to 6318)
  AUTHORS   Hartley,J.L. and Donelson,J.E.
  TITLE     Nucleotide sequence of the yeast plasmid
  JOURNAL   Nature 286, 860-865 (1980)
  MEDLINE   81012161
REFERENCE   3  (bases 3891 to 3990)
  AUTHORS   Broach,J.R., Guarascio,V.R. and Jayaram,M.
  TITLE     Recombination within the yeast plasmid 2-micron circle is
            site-specific
  JOURNAL   Cell 29, 227-234 (1982)
  MEDLINE   82259368
REFERENCE   4  (bases 3881 to 4020)
  AUTHORS   McLeod,M., Volkert,F. and Broach,J.R.
  TITLE     Components of the site-specific recombination system encoded by the
            yeast plasmid 2-micron circle
  JOURNAL   Cold Spring Harb. Symp. Quant. Biol. 49, 779-787 (1984)
  MEDLINE   85153059
REFERENCE   5  (bases 670 to 732)
  AUTHORS   Andrews,B.J., Proteau,G.A., Beatty,L.G. and Sadowski,P.D.
  TITLE     The FLP recombinase of the 2 micron circle DNA of yeast:
            Interaction with its target sequences
  JOURNAL   Cell 40, 795-803 (1985)
  MEDLINE   85176933
REFERENCE   6  (bases 5570 to 5605)
  AUTHORS   Babineau,D., Vetter,D., Andrews,B.J., Gronostajski,R.M.,
            Proteau,G.A., Beatty,L.G. and Sadowski,P.D.
  TITLE     The FLP protein of the 2-micron plasmid of yeast: Purification of
            the protein from Escherichia coli cells expressing the cloned FLP
            gene
  JOURNAL   J. Biol. Chem. 260, 12313-12319 (1985)
  MEDLINE   86008307
REFERENCE   7  (sites)
  AUTHORS   Gronostajski,R.M. and Sadowski,P.D.
  TITLE     Determination of DNA sequences essential for FLP-mediated
```



```
            recombination by a novel method
  JOURNAL   J. Biol. Chem. 260, 12320-12327 (1985)
  MEDLINE   86008308
REFERENCE   8  (sites)
  AUTHORS   Sutton,A. and Broach,J.R.
  TITLE     Signals for transcription initiation and termination in the
            Saccharomyces cerevisiae plasmid 2 micron circle
  JOURNAL   Mol. Cell. Biol. 5, 2770-2780 (1985)
  MEDLINE   86284639
REFERENCE   9  (sites)
  AUTHORS   Gronostajski,R.M. and Sadowski,P.D.
  TITLE     The FLP recombinase of the Saccharomyces cerevisiae 2-micron
            plasmid attaches covalently to DNA via a phosphotyrosyl linkage
  JOURNAL   Mol. Cell. Biol. 5, 3274-3279 (1985)
  MEDLINE   86310798
REFERENCE   10 (bases 667 to 739)
  AUTHORS   Senecoff,J.F., Bruckner,R.C. and Cox,M.M.
  TITLE     The FLP recombinase of the yeast 2-micron-m plasmid:
            Characterization of its recombination site
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82, 7270-7274 (1985)
  MEDLINE   86042647
REFERENCE   11 (sites)
  AUTHORS   McLeod,M., Craft,S. and Broach,J.R.
  TITLE     Identification of the crossover site during FLP-mediated
            recombination in the Saccharomyces cerevisiae plasmid 2 micron
            circle
  JOURNAL   Mol. Cell. Biol. 6, 3357-3367 (1986)
  MEDLINE   87089667
COMMENT     [8] sites; mRNA CAP sites and poly-adenylation sites. [9] sites;
            FLP binding.
            [7] sites; FLP cleavage.
            [11] sites; FLP-mediated recombination crossover site. Draft entry
            and clean copy sequence for [5] kindly provided by J.Senecoff,
            24-JAN-1986.
            Yeast 2 micron plasmid contains two 599 bp inverted repeats
            separated by a large unique (UL) and a small unique (US) region.
            During recombination the UL and US regions invert producing two
            sequence forms that differ in the orientation of one unique region
            relative to the other.  The A form is presented below. FLP is the
            only 2-micron circle-encoded protein needed for specific site
            recombination between the IRs of 2-micron circle.  The minimal size
            of the recombination site required for efficient FLP
            recombinase-catalyzed recombination in vitro is no more than 28 bp,
            which includes parts of two 13 bp inverted repeats (positions
            690-702 and 711-723) and all of an 8 bp spacer (703-710) [5].  The
            FLP recombinase cleaves the DNA at the boundaries of the spacer and
            becomes covalently linked to the spacer DNA [5],[9].  The
            efficiency of the recombination is reduced if the spacer in a
            recombinant site is increased or decreased by 1 bp, while the
            spacer in the second site is unaltered [5].  Recombination between
            two sites with identical 1-base pair additions or deletions is
            relatively unaffected, suggesting that pairing of sequences in the
            spacer regions is important in FLP-promoted recombination events
            [5].  The sequence asymmetry utilized by the recombinase to
            determine the orientation of the site is located uniquely within
            the spacer region. Another 13 bp direct repeat, is found at
            positions 676-688 [5].  FLP-mediated recombination involving two
            FLP sites that are inverted with respect to each other results in
            inversion of the DNA sequences between the sites [4].  If the
            participating recombination sites are in direct orientation, FLP
            promotes only the excision of the intervening DNA sequences [4].
            The Rep 1 and Rep proteins are involved plasmid partitioning and
            protein stability.
            A start codon in phase with the Repl coding region is located at
```

```
          positions 1966-1964.  Two CAP sites for Repl mRNA are located
          beyond the 'atg' codon (position 2008) at positions 2004 and 2005.
          Complete source information:
          Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71 [1],
          p82-6B [2], CV20 [3], pMMD2 [4], pGP20 [5], pJFS166 [10].

          NCBI gi: 172190
FEATURES             Location/Qualifiers
     source          1..6318
                     /organism="Saccharomyces cerevisiae"
     exon            1..545
     conflict        replace((157.160)..(157.160),"")
                     /citation=[1]
     conflict        replace((289.290)..(289.290),"")
                     /citation=[1]
     repeat_region   341..939
                     /note="IR2"
     conflict        replace((464.466)..(464.466),"")
                     /citation=[1]
     conflict        replace(558,"")
                     /citation=[1]
     conflict        replace(561,"")
                     /citation=[1]
     conflict        replace((622.624)..(622.624),"")
                     /citation=[1]
     conflict        replace(642,"")
                     /citation=[1]
     conflict        replace((665.666)..(665.666),"")
                     /citation=[1]
     misc_binding    673..722
                     /note="FLP recombinase binding site A [9]"
                     /bound_moiety="FLP recombinase"
     conflict        replace((793.794)..(793.794),"")
                     /citation=[1]
     mRNA            complement(836..2038)
                     /note="Repl mRNA (alt.)"
     mRNA            complement(836..2017)
                     /note="Repl mRNA (alt.)"
     mRNA            complement(836..2019)
                     /note="Repl mRNA (alt.)"
     mRNA            complement(836..2010)
                     /note="Repl mRNA (alt.)"
     mRNA            complement(836..2004)
                     /note="Repl mRNA (alt.)"
     mRNA            complement(836..2005)
                     /note="Repl mRNA (alt.)"
     CDS             complement(887..2008)
                     /note="Rep 1 protein;  NCBI gi: 172192"
                     /codon_start=1
                     /translation="MNGERLLACIKQCIMQHFQPMVYDESRCVIETTRGTFPVPDNYK
                     KYKTLAFAFVGHVLNTDDTPVIEKELDWPDPALVYNTIVDRIINHPELSQFISVAFIS
                     QLKATIGEGLDINVKGTLNRRGKGIRRPKGVFFRYMESPFVNTKVTAFFSYLRDYNKI
                     ASEYHNNTKFILTFSCQAYWASGPNFSALKNVIRCSIIHEYISKFVEREQDKGHIGDQ
                     ELPPEEDPSRELNNVQHEVNSLTEQDAEADEGLWGEIDSLCEKWQSEAEDQTEAEIIA
                     DRIIGNSQRMANLKIRRTKFKSVLYHILKELIQSQGTVKVYRGSSFSHDSIKISLHYE
                     EQHITAVWVYLTVKFEEHWKPVDVEVEFRCKFKERKVDG"
     mRNA            2254..2841
                     /note="D mRNA (alt.; 5' end +/- 3 bp)"
     mRNA            2254..2861
                     /note="D mRNA (alt.; 5' end +/- 3 bp)"
     CDS             2271..2816
                     /note="D protein;  NCBI gi: 172193"
                     /codon_start=1
                     /translation="MPYKTAIDCIEELATQCFLSKLTDDDVSTFRRVCSKENDIIKLA
```



```
                     LRIPRTIDYTSILRLLYDTLPLRSLSFNEALPLFCYSIDPAQQRQCDLRFYLRDVVKL
                     ARPRKRLEMQKALLQWLPSLLSDVTLQLLNDIRIRFEEIQPNIRQTVLQIYDRTCYPS
                     LNFEHPNLGVFPETDSIFEPV"
     repeat_region   3714..4312
                     /note="IR1"
     misc_binding    3930..3979
                     /note="FLP-recombinase binding site B [9]"
                     /bound_moiety="FLP recombinase"
     mRNA            complement(4108..5182)
                     /note="REP2 mRNA (major alt.)"
     mRNA            complement(4108..5183)
                     /note="REP2 mRNA (major alt.)"
     mRNA            complement(4108..5184)
                     /note="REP2 mRNA (major alt.)"
     mRNA            complement(4108..5223)
                     /note="REP2 mRNA (minor alt.)"
     mRNA            complement(4108..5195)
                     /note="REP2 mRNA (major alt.)"
...
Note: remainder of annotations omitted.

  Query Match          100.0%;  Score 26;  DB 43;  Length 6318;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3930 gaagttcctatactttctagagaataggaacttc 3963
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   20
LOCUS       SCA21        6318 bp    DNA   circular  PLN       29-JUN-1995
DEFINITION  2 micron plasmid of yeast (circularly closed).
ACCESSION   V01323 J01347 L00321 L00322 L00323 L00324 M10185 M11111 M11593
            M14239 M14240 M14241 M14242 M14243 M14244 M14245 M14253 M14254
            M14255 M14256 M14257 M14258 M14259 M14591 M14592 M14593 M14594
            M14596 M14597 M14598
KEYWORDS    circular; origin of replication.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryotae; mitochondrial eukaryotes; Eumycota; Ascomycotina;
            Hemiascomycetes; Saccharomycetales; Saccharomycetaceae;
            Saccharomyces.
REFERENCE   1  (bases 1 to 6318)
  AUTHORS   Hartley,J.L. and Donelson,J.E.
  TITLE     Nucleotide sequence of the yeast plasmid
  JOURNAL   Nature 2869 , 860-865 (1980)
  MEDLINE   81012161
COMMENT     NCBI gi: 4182
FEATURES             Location/Qualifiers
     source          1..6318
                     /organism="Saccharomyces cerevisiae"
                     /plasmid="2 micron plasmid"
     CDS             complement(887..2008)
                     /note="NCBI gi: 4183"
                     /codon_start=1
                     /product="protein Baker"
                     /translation="MNGERLLACIKQCIMQHFQPMVYDESRCVIETTRGTFPVPDNYK
                     KYKTLAFAFVGHVLNTDDTPVIEKELDWPDPALVYNTIVDRIINHPELSQFISVAFIS
                     QLKATIGEGLDINVKGTLNRRGKGIRRPKGVFFRYMESPFVNTKVTAFFSYLRDYNKI
                     ASEYHNNTKFILTFSCQAYWASGPNFSALKNVIRCSIIHEYISKFVEREQDKGHIGDQ
                     ELPPEEDPSRELNNVQHEVNSLTEQDAEADEGLWGEIDSLCEKWQSEAEDQTEAEIIA
                     DRIIGNSQRMANLKIRRTKFKSVLYHILKELIQSQGTVKVYRGSSFSHDSIKISLHYE
                     EQHITAVWVYLTVKFEEHWKPVDVEVEFRCKFKERKVDG"
```

```
     CDS             complement(4308..5198)
                     /note="NCBI gi: 4184"
                     /codon_start=1
                     /product="protein Charlie"
                     /translation="MDDIETAKNLTVKARTAYSVWDVCRLFIEMIAPDVDIDIESKRK
                     SDELLFPGYVIRPMESLTTGRPYGLDSSAEDSSVSSDSSAEVILPAAKMVKERFDSIG
                     NGMLSSQEASQAAIDLMLQNNKLLDNRKQLYKSIAIIIGRLPEKDKKRATEMLMRKMD
                     CTQLLVPPAPTEEDVMKLVSVVTQLLTLVPPDRQAALIGDLFIPESLKDIFNSFNELA
                     AENRLQQKKSELEGRTEVNHANTNEEVPSRRTRSRDTNARGAYKLQNTITEGPKAVPT
                     KKRRVATRVRGRKSRNTSRV"
BASE COUNT     1876 a   1284 c   1179 g   1979 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 41;  Length 6318;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3930 gaagttcctatactttctagagaataggaacttc 3963
         ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   21
LOCUS       SYNECOYST    6445 bp    DNA             SYN       27-JUL-1994
DEFINITION  Cloning vector sequence (E. coli/yeast/phage P1) ADH2 gene,
            promoter; beta-lactamase gene; TRP1 gene; CYC1 gene, terminator;
            and two replication origins.
ACCESSION   L11060
KEYWORDS    .
SOURCE      Cloning vector DNA.
  ORGANISM  Cloning vector
            Artificial sequences; Cloning vehicles.
REFERENCE   1  (sites)
  AUTHORS   Brunelli,J.P. and Pall,M.L.
  TITLE     A series of yeast shuttle vectors for expression of cDNAs and other
            DNA sequences
  JOURNAL   Yeast 9, 1299-1308 (1993)
  MEDLINE   94205259
REFERENCE   2  (bases 1 to 6445)
  AUTHORS   Brunelli,J.P. and Pall,M.L.
  TITLE     A series of yeast/Escherichia coli lambda expression vectors
            designed for directional cloning of cDNAs and cre/lox-mediated
            plasmid excision
  JOURNAL   Yeast 9, 1309-1318 (1993)
  MEDLINE   94205260
REFERENCE   3  (bases 1 to 6445)
  AUTHORS   Pall,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-1993) Martin L. Pall, Department of Genetics and
            Cell Biology, Washington State University, Pullman, WA 99164-4234,
            USA
COMMENT     NCBI gi: 310741
FEATURES             Location/Qualifiers
     source          1..6445
                     /organism="Cloning vector"
                     /sequenced_mol="DNA"
     misc_feature    1..80
                     /standard_name="polylinker"
     promoter        81..668
                     /gene="ADH2"
     rep_origin      750..1500
     misc_feature    1799..2889
                     /note="beta-lactamase"
     misc_feature    2890..3790
```



```
                     /note="TRP1 gene"
     rep_origin      3790..5465
                     /note="2 micron origin"
     misc_feature    5660..5745
                     /note="lox site"
     misc_feature    5785..5850
                     /note="lox site"
     terminator      5880..6420
                     /standard_name="CYC1 terminator"
BASE COUNT     1790 a   1354 c   1443 g   1858 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 6445;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5068 gaagttcctatactttctagagaataggaacttc 5101
         ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   22
ID   CV37458    standard; circular DNA; FUN; 6624 BP.
AC   U37458;
DT   10-NOV-1995 (Rel. 45, Created)
DT   10-NOV-1995 (Rel. 45, Last updated, Version 1)
DE   Yeast CUP1 expression-multicopy (2micron) cloning vector YRTAG300
DE   with the hemagglutinin tag sequence, complete sequence. selection).
KW   .
OS   Saccharomyces cerevisiae (yeast)
OC   Eukaryota; Plantae; Thallobionta; Eumycota; Hemiascomycetes;
OC   Endomycetales; Saccharomycetaceae.
RN   [1]
RP   1-6624
RA   Lieberman B.;
RT   ;
RL   Submitted (03-OCT-1995) to the EMBL/GenBank/DDBJ databases.
RL   Benjamin Lieberman, Pharmacology, Duke University, Research Drive,
RL   P.O. 3813, Durham, NC 27710, USA
CC   NCBI gi: 1052969
FH   Key             Location/Qualifiers
FH
FT   source          1..6624
FT                   /organism="Saccharomyces cerevisiae"
FT                   /note="based on pRS424 (TRP selection); includes CUP1
FT                   promoter, Cyc terminator and the hemagglutinin coding
FT                   region fused in frame to a start codon and and two
FT                   restriction sites (SstI and XhoI)"
FT   CDS             574..1434
FT                   /note="NCBI gi: 1052970"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /product="beta-lactamase"
FT                   /note="pid:g1052970"
FT   misc_feature    3040..3098
FT                   /note="HA TAG with start codon (EcoRI-HATAG-SSTI-XHOI)"
FT   CDS             complement(4514..5215)
FT                   /gene="TRP1"
FT                   /note="NCBI gi: 1052971"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /product="N-(5'-phosphoribosyl)-anthranilate isomerase"
FT                   /note="pid:g1052971"
SQ   Sequence 6624 BP; 1845 A; 1543 C; 1397 G; 1839 T; 0 other;
```

```
  Query Match          100.0%;  Score 26;  DB 2;  Length 6624;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   6070 gaagttcctattctctagaaagtataggaacttc 6103
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   23
LOCUS       CVU37458     6624 bp    DNA   circular  SYN       08-NOV-1995
DEFINITION  Yeast CUP1 expression-multicopy (2micron) cloning vector YRTAG300
            with the hemagglutinin tag sequence, complete sequence. selection).
ACCESSION   U37458
KEYWORDS    .
SOURCE      Cloning vector YRTAG300.
  ORGANISM  Cloning vector YRTAG300
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 6624)
  AUTHORS   Lieberman,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Benjamin Lieberman, Pharmacology, Duke
            University, Research Drive, P.O. 3813, Durham, NC 27710, USA
COMMENT     NCBI gi: 1052969
FEATURES             Location/Qualifiers
     source          1..6624
                     /organism="Saccharomyces cerevisiae"
                     /note="based on pRS424 (TRP selection); includes CUP1
                     promoter, Cyc terminator and the hemagglutinin coding
                     region fused in frame to a start codon and and two
                     restriction sites (SstI and XhoI)"
     CDS             574..1434
                     /note="NCBI gi: 1052970"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     misc_feature    3040..3098
                     /note="HA TAG with start codon (EcoRI-HATAG-SSTI-XHOI)"
     CDS             complement(4514..5215)
                     /gene="TRP1"
                     /note="NCBI gi: 1052971"
                     /codon_start=1
                     /transl_table=11
                     /product="N-(5'-phosphoribosyl)-anthranilate isomerase"
                     /translation="MKHTKAAWSMSVINFTGSSGPLVKVCGLQSTEAAECALDSDADL
                     LGIICVPNRKRTIDPVIARKISSLVKAYKNSSGTPKYLVGVFRNQPKEDVLALVNDYG
                     IDIVQLHGDESWQEYQEFLGLPVIKRLVFPKDCNILLSAASQKPHSFIPLFDSEAGGT
                     GELLDWNSISDWVGRQESPESLHFMLAGGLTPENVGDALRLNGVIGVDVSGGVETNGV
                     KDSNKIANFVKNAKK"
BASE COUNT     1845 a   1543 c   1397 g   1839 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 84;  Length 6624;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   6070 gaagttcctattctctagaaagtataggaacttc 6103
```



```
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   24
LOCUS       PRS425       6849 bp    DNA   circular  SYN       24-MAY-1995
DEFINITION  Yeast episomal vector pRS425 with LEU2 marker, complete sequence.
ACCESSION   U03452
KEYWORDS    .
SOURCE      Cloning vector pRS425.
  ORGANISM  Cloning vector pRS425
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 6849)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 6849)
  AUTHORS   Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and
            Hieter,P.
  TITLE     Multifunctional yeast high-copy-number shuttle vectors
  JOURNAL   Gene 110 (1), 119-122 (1992)
  MEDLINE   92184105
REFERENCE   3  (bases 1 to 6849)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416323
FEATURES             Location/Qualifiers
     source          1..6849
                     /organism="cloning vector pRS425"
BASE COUNT     1869 a   1504 c   1543 g   1933 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 6849;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   6402 gaagttcctatactttctagagaataggaacttc 6435
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   25
LOCUS       SYNYEP24V    7769 bp    DNA   circular  SYN       07-JUL-1993
DEFINITION  YEp24 yeast extrachromosomal plasmid.
ACCESSION   L09156
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  Synthetic construct
            Artificial sequences.
REFERENCE   1  (bases 1 to 7769)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program. YEp24 -
            Yeast Extrachromosomal plasmid
            ENTRY YEP24                    #TYPE DNA CIRCULAR
            TITLE YEp24 - Yeast Extrachromosomal plasmid
            DATE      12-SEP-1986
```

```
            #sequence  16-DEC-1986
          ACCESSION VB0067
          SOURCE    artificial
          REFERENCE
            #number 1
            #authors  Botstein D., Falco S.C., Stewart S.E., Brennan M.,
                      Scherer S., Stinchcomb D.T., Struhl K., Davis R.W.
            #journal Gene (1979) 8: 17-24
          REFERENCE
            #number 2
            #citation sequence information from BioLabs
          REFERENCE
            #number 3
            #authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
            #book    Cloning Vectors, Elsvier 1985 and supplements
            #comment  vector VI-A-i-5
          COMMENT
            Obtained 12-SEP-1986 from New England Biolabs
            on magnetic tape
            Revised 16-DEC-1986 by F. Pfeiffer:
            6140/1 'AT' to 'TA' to match revised sequence of PBR322 COMMENT
            The tetracycline resistance promoter was separated from coding
            sequence by the URA3 gene. Tc-R is dependent on the construct.
          KEYWORDS
            URA3 = orotidine-5'-phosphate decarboxylase (EC 4.1.1.23)
          CROSSREFERENCE
            #parent
              VecBase(3):pBR322, GenBank(50):YSCPlasm, GenBank(50):YSCODCD
          PARENT
            Features of YEp24 (7769 bp)
              residue      source
                1- 340     1- 340     2u-plasmid
              341- 939   341- 939     2u-plasmid
              341- 939  4312-3714 (c) 2u-plasmid
              940-2247  3713-2406 (c) 2u-plasmid
            2244-2278     1-  35     pBR322
            2273-3442     1-1170     URA3 gene (GenBank(50):YSCODCD)
            3438-7769    29-4360     pBR322
            Conflict (cfl) and Mutations (mut): none
          FEATURE
            3495-4685   1-1191     Tc-R; tetracycline resistance protein
            6707-7495 789-   1 (c) Ap-R; b-lactamase
            2499-3302 227-1030     URA3 gene from S. cerevisiae +D4
          POLYLINKER
          SELECTION
            #resistance Ap, Tc
            #enzyme  orotidine-5'-phosphate decarboxylase (URA3) SUMMARY
          YEp24    #length 7769   #checksum 7274.

          NCBI gi: 310855
FEATURES             Location/Qualifiers
     source          1..7769
                     /organism="Synthetic construct"
                     /sequenced_mol="DNA"
BASE COUNT     2102 a   1830 c   1844 g   1993 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 7769;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    690 gaagttcctattctctagaaagtataggaacttc 723
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34
```



```
RESULT   26
LOCUS       CVPFL46L     7822 bp    DNA             SYN       15-AUG-1995
DEFINITION  multicopy Saccharomyces cerevisiae/E. coli shuttle vector.
ACCESSION   X70269
KEYWORDS    2-micron yeast replication origin; LEU2 selectable marker;
            pUC19 plasmid.
SOURCE      cloning vectors.
  ORGANISM  cloning vectors
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 644; 3499 to 3625; 5877 to 7822)
  AUTHORS   Yanisch-Perron,C., Vieira,J. and Messing,J.
  TITLE     Improved M13 phage cloning vectors and host strains: nucleotide
            sequences of the M13mp18 and pUC19 vectors
  JOURNAL   Gene 33 (1), 103-119 (1985)
  MEDLINE   85180545
REFERENCE   2  (bases 644 to 3499)
  AUTHORS   Struhl,K., Stinchcomb,D.T., Scherer,S. and Davis,R.W.
  TITLE     High-frequency transformation of yeast: autonomous replication of
            hybrid DNA molecules
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 76 (3), 1035-1039 (1979)
  MEDLINE   79180126
REFERENCE   3  (bases 3625 to 5877)
  AUTHORS   Chevalier,M.R. and Lacroute,F.
  TITLE     Trancriptional and traductional expression of a chimeric
            bacterial- yeast plasmid in yeast
  JOURNAL   Gene 1, 11-19 (1980)
REFERENCE   4  (bases 1 to 7822)
  AUTHORS   Ozier-Kalogeropoullos,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-1993) to the EMBL/GenBank/DDBJ databases.
            Ozier-Kalogeropoullos O., CGM, CNRS, 91190 Gif sur Yvette, France
            e-mail:odile%FRCGM51.BITNET@vm.gmd.de
REFERENCE   5  (bases 1 to 7822)
  AUTHORS   Bonneaud,N., Ozier-Kalogeropoulos,O., Li,G.Y., Labouesse,M.,
            Minvielle-Sebastia,L. and Lacroute,F.
  TITLE     A family of low and high copy replicative, integrative and
            single-stranded S. cerevisiae/E. coli shuttle vectors
  JOURNAL   Yeast 7 (6), 609-615 (1991)
  MEDLINE   92116645
COMMENT     The pFL46L was constructed from pUC19 plasmid where two alu I sites
            were modified.The site 629 was replaced by a BglII linker and the
            site 747 by a ClaI site. The yeast selectable marker has been
            cloned in the BglII site and the 2 micron 2.2 kb EcoRI fragment
            containing ORI and STB gene has been cloned at the ClaI site. The
            pFL46L is described in Bonneaud et al (1991): 'A family of low and
            high copy replicative, integrative and single-stranded
            S.cerevisiae/E.coli shuttle vectors'. YEAST, 7, 609-615.

            NCBI gi: 397134
FEATURES             Location/Qualifiers
     source          1..7822
                     /organism="Cloning vector"
BASE COUNT     2222 a   1664 c   1691 g   2245 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 7822;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5154 gaagttcctatactttctagagaataggaacttc 5187
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1
```

```
RESULT   27
ID   CV33753    standard; circular DNA; SYN; 7834 BP.
AC   U33753;
DT   12-OCT-1995 (Rel. 45, Created)
DT   12-OCT-1995 (Rel. 45, Last updated, Version 1)
DE   Yeast episomal cloning vector pADNS, with ADH1 promoter, complete
DE   sequence.
KW   .
OS   Cloning vector pADNS
OC   Artificial sequences; Cloning vectors.
RN   [1]
RP   1-7834
RX   MEDLINE; 89264471.
RA   Colicelli J., Birchmeier C., Michaeli T., O'Neill K., Riggs M.,
RA   Wigler M.;
RT   "Isolation and characterization of a mammalian gene encoding a
RT   high-affinity cAMP phosphodiesterase";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:3599-3603(1989).
RN   [2]
RP   1-7834
RA   Stillman D.J.;
RT   ;
RL   Submitted (10-AUG-1995) to the EMBL/GenBank/DDBJ databases.
RL   David J. Stillman, Division of Molecular Biology and Genetics,
RL   Department of Oncological Sciences, University of Utah Heath
RL   Sciences Center, Salt Lake City, UT 84132, USA
CC   NCBI gi: 988314
FH   Key             Location/Qualifiers
FH
FT   source          1..7834
FT                   /organism="Cloning vector pADNS"
FT   misc_feature    1..38
FT                   /note="polylinker"
FT   terminator      39..505
FT                   /gene="ADH1"
FT   promoter        6381..7834
FT                   /gene="ADH1"
SQ   Sequence 7834 BP; 2192 A; 1672 C; 1646 G; 2324 T; 0 other;

  Query Match          100.0%;  Score 26;  DB 15;  Length 7834;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5763 gaagttcctatactttctagagaataggaacttc 5796
         ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   28
LOCUS       CVU33753     7834 bp    DNA   circular  SYN       19-SEP-1995
DEFINITION  Yeast episomal cloning vector pADNS, with ADH1 promoter, complete
            sequence.
ACCESSION   U33753
KEYWORDS    .
SOURCE      cloning vectors.
  ORGANISM  Cloning vector pADNS
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 7834)
  AUTHORS   Colicelli,J., Birchmeier,C., Michaeli,T., O'Neill,K., Riggs,M. and
            Wigler,M.
  TITLE     Isolation and characterization of a mammalian gene encoding a
            high-affinity cAMP phosphodiesterase
```



```
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 86 (10), 3599-3603 (1989)
  MEDLINE   89264471
REFERENCE   2  (bases 1 to 7834)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-1995) David J. Stillman, Division of Molecular
            Biology and Genetics, Department of Oncological Sciences,
            University of Utah Heath Sciences Center, Salt Lake City, UT 84132,
            USA
COMMENT     NCBI gi: 988314
FEATURES             Location/Qualifiers
     source          1..7834
                     /organism="cloning vector pADNS"
     misc_feature    1..38
                     /note="polylinker"
     terminator      39..505
                     /gene="ADH1"
     promoter        6381..7834
                     /gene="ADH1"
BASE COUNT     2192 a   1672 c   1646 g   2324 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 7834;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   5763 gaagttcctatactttctagagaataggaacttc 5796
         ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   29
LOCUS       A17115       7859 bp    DNA             PAT       31-MAR-1994
DEFINITION  Yeast expression vector pSW from S.cerevisiae (SEQ ID NO: 15).
ACCESSION   A17115
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 7859)
  AUTHORS   .
  TITLE     STEM CELL INHIBITING PROTEINS
  JOURNAL   Patent: WO 9313206-A 15 08-JUL-1993;

COMMENT     NCBI gi: 512887
FEATURES             Location/Qualifiers
     source          1..7859
                     /organism="Artificial sequences"
BASE COUNT     2317 a   1656 c   1600 g   2286 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 34;  Length 7859;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
         ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   30
LOCUS       I13185       7859 bp    DNA             PAT       19-JUL-1995
DEFINITION  Sequence 4 from patent US 5434073.
ACCESSION   I13185
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7859)
  AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
  TITLE     Fibrinolytic and anti-thrombotic cleavable dimers
  JOURNAL   Patent: US 5434073-A 4 18-JUL-1995;

COMMENT     NCBI gi: 910533
FEATURES             Location/Qualifiers
     source          1..7859
                     /organism="unknown"
BASE COUNT     2317 a   1656 c   1600 g   2286 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 35;  Length 7859;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
         ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   31
ID   A19996     standard; DNA; SYN; 7859 BP.
AC   A19996;
DT   14-JUL-1995 (Rel. 44, Created)
DT   14-JUL-1995 (Rel. 44, Last updated, Version 1)
DE   SEQ ID NO: 4; Synthetic plasmid pSW6.
KW   .
OS   None
OC   Artificial sequences.
RN   [1]
RA   ;
RT   "PROTEINS AND NUCLEIC ACIDS";
RL   Patent number WO9109125-A/4, 27-JUN-1991.
FH   Key             Location/Qualifiers
FH
FT   source          1..7859
FT                   /organism="Artificial sequences"
SQ   Sequence 7859 BP; 2317 A; 1656 C; 1600 G; 2286 T; 0 other;

  Query Match          100.0%;  Score 26;  DB 88;  Length 7859;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
         ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   32
LOCUS       A18079       7984 bp    DNA             PAT       22-APR-1994
DEFINITION  yeast expression vector pSW6 seq ID No: 19.
ACCESSION   A18079
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 7984)
  AUTHORS   .
  TITLE     PHARMACEUTICALLY ACTIVE PROTEINS COMPRISING AN ACTIVE PROTEIN AND
```



```
            AN INTEGRINAFFINITY SEQUENCE
  JOURNAL   Patent: WO 9207874-A 33 14-MAY-1992;

COMMENT     NCBI gi: 513171
FEATURES             Location/Qualifiers
     source          1..7984
                     /organism="Artificial sequences"
                     /lab_host="yeast expression vector pSW6"
BASE COUNT     2345 a   1695 c   1638 g   2306 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 34;  Length 7984;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
         ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   33
LOCUS       CVU29899     8117 bp    DNA   circular  SYN       01-AUG-1995
DEFINITION  Cloning vector pACT2 MatchmakerII, complete sequence.
ACCESSION   U29899
KEYWORDS    .
SOURCE      Cloning vector pACT2.
  ORGANISM  Cloning vector pACT2
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 8117)
  AUTHORS   Kitts,P.A.
  TITLE     Clontech Vectors On Disk, version 1.3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 8117)
  AUTHORS   Stile,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-1995) John S. Stile, Clontech Laboratories, Inc.,
            4030 Fabian Way, Palo Alto, CA 94303, USA
COMMENT     This vector can be obtained from CLONTECH Laboratories, Inc., 4030
            Fabian  Way, Palo Alto, CA 94303, USA. To place an order call
            (415) 424-8222 or  (800) 662-2566, extension 1. International
            customers, please contact your  local distributor. For technical
            information, call (415) 424-8222 or  (800)  662-2566, extension 3.
            This sequence has been compiled from information in the sequence
            databases,  published literature and other sources, together with
            partial sequences  obtained by CLONTECH. If you suspect there is
            an error in this sequence,  please contact CLONTECH's Technical
            Service Department at (415) 424-8222  or (800) 662-2566, extension
            3 or E-mail TECH@CLONTECH.COM.

            NCBI gi: 915409
FEATURES             Location/Qualifiers
     source          1..8117
                     /organism="Cloning vector pACT2"
     rep_origin      1..2057
                     /note="Yeast 2 micron ori"
     CDS             2474..3568
                     /note="NCBI gi: 915410"
                     /codon_start=1
                     /transl_table=11
                     /product="LEU2"
                     /translation="MSAPKKIVVLPGDHVGQEITAEAIKVLKAISDVRSNVKFDFENH
                     LIGGAAIDATGVPLPDEALEASKKVDAVLLGAVGGPKWGTGSVRPEQGLLKIRKELQL
                     YANLRPCNFASDSLLDLSPIKPQFAKGTDFVVVRELVGGIYFGKRKEDDGDGVAWDSE
                     QYTVPEVQRITRMAAFMALQHEPPLPIWSLDKANVLASSRLWRKTVEETIKNEFPTLK
```

```
                    VQHQLIDSAAMILVKNPTHLNGIIITSNMFGDIISDEASVIPGSLGLLPSASLASLPD
                    KNTAFGLYEPCHGSAPDLPKNKVDPIATILSAAMMLKLSLNLPEEGKAIEDAVKKVLD
                    AGIRTGDLGGSNSTTEVGDAVAEEVKKILA"
     misc_feature   group(4268..4327,4367..4412)
                    /note="Lox Sites"
     terminator     4415..4742
                    /note="S. cerevisiae ADH1 terminator"
     misc_feature   5042..5068
                    /note="encodes HA epitope which fuses to Gal 4"
     CDS            complement(<5081..5488)
                    /note="encodes 5' region of the Gal 4 activation domain
                    polypeptide; stop codon is located downstream within the
                    multiple cloning site;  NCBI gi: 915411"
                    /codon_start=1
                    /transl_table=11
                    /product="Gal 4 activation domain polypeptide"
                    /translation="MDKAELIPEPPKKKRKVELGTAANFNQSGNIADSSLSFTFTNSS
                    NGPNLITTQTNSQALSQPIASSNVHDNFMNNEITASKIDDGNNSKPLSPGWTDQTAYN
                    AFGITTGMFNTTTMDDVYNYLFDDEDTPPNPKKE"
     promoter       5504..5901
                    /note="S. cerevisiae ADH1 Promoter"
     rep_origin     6336..6979
                    /note="pBR322 origin of replication"
     CDS            complement(7127..7987)
                    /note="NCBI gi: 915412"
                    /codon_start=1
                    /transl_table=11
                    /product="beta-lactamase"
                    /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                    IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRVDAGQEQLGRRIHYSQNDLVE
                    YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                    DRWEPELNEAIPNDERDTTMPAAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                    LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                    EIGASLIKHW"
     -10_signal     8029..>8034
     -35_signal     8052..8057
BASE COUNT     2423 a   1677 c   1757 g   2260 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 8117;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    584 gaagttcctattctctagaaagtataggaacttc 617
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   34
ID   CV30497    standard; circular DNA; SYN; 8393 BP.
AC   U30497;
DT   12-OCT-1995 (Rel. 45, Created)
DT   12-OCT-1995 (Rel. 45, Last updated, Version 1)
DE   Cloning vector pAS2-1, complete sequence.
KW   .
OS   Cloning vector pAS2-1
OC   Artificial sequences; Cloning vectors.
RN   [1]
RP   1-8393
RA   Kitts P.A.;
RT   "CLONTECH Vectors On Disk, version 1.3";
RL   Unpublished.
RN   [2]
RP   1-8393
```



```
RA   Stile J.S.;
RT   ;
RL   Submitted (28-JUN-1995) to the EMBL/GenBank/DDBJ databases.
RL   John S. Stile, CLONTECH Laboratories, Inc., 4030 Fabian Way, Palo
RL   Alto, CA 94303, USA
CC   This vector can be obtained from CLONTECH Laboratories, Inc., 4030
CC   Fabian Way, Palo Alto, CA 94303, USA. To place an order call (415)
CC   424-8222 or (800) 662-2566, extension 1. International customers,
CC   please contact your local distributor. For technical information,
CC   call (415) 424-8222 or (800) 662-2566, extension 3. This sequence
CC   has been compiled from information in the sequence databases,
CC   published literature and other sources, together with partial
CC   sequences obtained by CLONTECH. If you suspect there is an error in
CC   this sequence, please contact CLONTECH's Technical Service
CC   Department at (415) 424-8222 or (800) 662-2566, extension 3 or
CC   E-mail TECH@CLONTECH.COM. NCBI gi: 988208
FH   Key             Location/Qualifiers
FH
FT   source          1..8393
FT                   /organism="Cloning vector pAS2-1"
FT   rep_origin      1..1348
FT                   /note="two micron origin of replication (B form)"
FT   CDS             1884..2558
FT                   /gene="TRP1"
FT                   /note="NCBI gi: 988209"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /note="pid:g988209"
FT   rep_origin      2609..3277
FT                   /note="f1+ origin"
FT   CDS             complement(4018..4305)
FT                   /gene="CYH2"
FT                   /note="NCBI gi: 988210"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /note="pid:g988210"
FT   promoter        4768..5475
FT                   /note="from ADH1 gene"
FT   CDS             5502..6065
FT                   /note="contains Gal4 binding domain and epitopes for
FT                   monoclonal antibody binding; NCBI gi: 988211"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /product="fusion protein"
FT                   /note="pid:g988211"
FT   misc_feature    5502..5942
FT                   /note="encodes Gal4 binding domain"
FT   misc_feature    5953..5970
FT                   /note="encodes epitope for monoclonals D11 and F10 binding
     "
FT   misc_feature    5971..6016
FT                   /note="multiple cloning site"
FT   terminator      6032..6224
FT                   /note="from ADH1 gene; contains stop codons for frame 1,2
FT                   and 3"
FT   rep_origin      6232..7403
FT                   /note="pUC origin of replication"
FT   CDS             complement(7403..8263)
FT                   /gene="ampicillin resistance"
FT                   /note="NCBI gi: 988212"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /note="pid:g988212"
SQ   Sequence 8393 BP; 2351 A; 1754 C; 1871 G; 2416 T; 1 other;
```

```
  Query Match          100.0%;  Score 26;  DB 15;  Length 8393;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    965 gaagttcctatactttctagagaataggaacttc 998
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   35
LOCUS       CVU30497     8393 bp    DNA   circular  SYN       19-SEP-1995
DEFINITION  Cloning vector pAS2-1, complete sequence.
ACCESSION   U30497
KEYWORDS    .
SOURCE      cloning vectors.
  ORGANISM  Cloning vector pAS2-1
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 8393)
  AUTHORS   Kitts,P.A.
  TITLE     CLONTECH Vectors On Disk, version 1.3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 8393)
  AUTHORS   Stile,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-1995) John S. Stile, CLONTECH Laboratories, Inc.,
            4030 Fabian Way, Palo Alto, CA 94303, USA
COMMENT     This vector can be obtained from CLONTECH Laboratories, Inc.,  4030
            Fabian  Way, Palo Alto, CA 94303, USA.  To place an order call
            (415) 424-8222 or (800) 662-2566, extension 1.  International
            customers, please contact your local distributor.  For technical
            information, call (415) 424-8222 or (800)  662-2566, extension 3.
            This sequence has been compiled from information in the sequence
            databases,  published literature and other sources, together with
            partial sequences obtained by CLONTECH. If you suspect there is an
            error in this sequence,  please contact CLONTECH's Technical
            Service Department at (415) 424-8222 or (800) 662-2566, extension 3
            or E-mail TECH@CLONTECH.COM.

            NCBI gi: 988208
FEATURES             Location/Qualifiers
     source          1..8393
                     /organism="Cloning vector pAS2-1"
     rep_origin      1..1348
                     /note="two micron origin of replication (B form)"
     CDS             1884..2558
                     /gene="TRP1"
                     /note="NCBI gi: 988209"
                     /codon_start=1
                     /transl_table=11
                     /translation="MSVINFTGSSGPLVKVCGLQSTEAAECALDSDADLLGIICVPNR
                     KRTIDPVIARKISSLVKAYKNSSGTPKYLVGVFRNQPKEDVLALVNDYGIDIVQLHGD
                     ESWQEYQEFLGLPVIKRLVFPKDCNILLSAASQKPHSFIPLFDSEAGGTGELLDWNSI
                     SDWVGRQESPESLHFMLAGGLTPENVGDALRLNGVIGVDVSGGVETNGVKDSNKIANF
                     VKNAKK"
     rep_origin      2609..3277
                     /note="f1+ origin"
     CDS             complement(4018..4305)
                     /gene="CYH2"
                     /note="NCBI gi: 988210"
                     /codon_start=1
                     /transl_table=11
                     /translation="MPSRFTKTRKHRGHVSAGKGRIGKHRKHPGGRGMAGGQHHHRIN
                     MDKYHPGYFGKXWYEILPQATSSFLEASLELGQIVDIDPRRQERPILEICF"
```



```
     promoter        4768..5475
                     /note="from ADH1 gene"
     CDS             5502..6065
                     /note="contains Gal4 binding domain and epitopes for
                     monoclonal antibody binding;  NCBI gi: 988211"
                     /codon_start=1
                     /transl_table=11
                     /product="fusion protein"
                     /translation="MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKT
                     KRSPLTRAHLTEVESRLERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNV
                     NKDAVTDRLASVETDMPLTLRQHRISATSSSEESSNKGQRQLTVSPELQYPALTHMAM
                     EAEFPGIRRPAAKLIPGEFLMIYDFYY"
     misc_feature    5502..5942
                     /note="encodes Gal4 binding domain"
     misc_feature    5953..5970
                     /note="encodes epitope for monoclonals D11 and F10
                     binding"
     misc_feature    5971..6016
                     /note="multiple cloning site"
     terminator      6032..6224
                     /note="from ADH1 gene; contains stop codons for frame 1,2
                     and 3"
     rep_origin      6232..7403
                     /note="pUC origin of replication"
     CDS             complement(7403..8263)
                     /gene="ampicillin resistance"
                     /note="NCBI gi: 988212"
                     /codon_start=1
                     /transl_table=11
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT     2351 a   1754 c   1871 g   2416 t      1 others
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 8393;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    965 gaagttcctatactttctagagaataggaacttc 998
          ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   36
LOCUS       YEP213      10667 bp    DNA   circular  SYN       17-NOV-1993
DEFINITION  Yeast episomal vector YEp213, complete sequence.
ACCESSION   U03499
KEYWORDS    .
SOURCE      Cloning vector YEp213.
  ORGANISM  Cloning vector YEp213
            Artificial sequences; Cloning vector.
REFERENCE   1  (bases 1 to 10667)
  AUTHORS   Rose,A.B. and Broach,J.R.
  TITLE     Propagation and expression of cloned genes in yeast: 2-umcircle
            based vectors
  JOURNAL   Meth. Enzymol. 185, 234-279 (1990)
  MEDLINE   90340124
REFERENCE   2  (bases 1 to 10667)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (16-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416341
FEATURES             Location/Qualifiers
     source          1..10667
                     /lab_host="Saccharomyces cerevisiae"
                     /organism="Cloning vector YEp213"
BASE COUNT     2837 a   2459 c   2350 g   3021 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 10667;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   9951 gaagttcctatactttctagagaataggaacttc 9984
         ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   37
LOCUS       YEP13       10667 bp    DNA   circular  SYN       17-NOV-1993
DEFINITION  Yeast episomal vector YEp13, complete sequence.
ACCESSION   U03498
KEYWORDS    .
SOURCE      Cloning vector YEp13.
  ORGANISM  Cloning vector YEp13
            Artificial sequences; Cloning vector.
REFERENCE   1  (bases 1 to 10667)
  AUTHORS   Rose,A.B. and Broach,J.R.
  TITLE     Propagation and expression of cloned genes in yeast: 2-umcircle
            based vectors
  JOURNAL   Meth. Enzymol. 185, 234-279 (1990)
  MEDLINE   90340124
REFERENCE   2  (bases 1 to 10667)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416340
FEATURES             Location/Qualifiers
     source          1..10667
                     /lab_host="Saccharomyces cerevisiae"
                     /organism="Cloning vector YEp13"
BASE COUNT     2889 a   2443 c   2366 g   2969 t
ORIGIN

  Query Match          100.0%;  Score 26;  DB 61;  Length 10667;
  Best Local Similarity  76.5%;  Pred. No. 4.89e-08;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   9951 gaagttcctatactttctagagaataggaacttc 9984
         ||||||||||||||        ||||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   38
LOCUS       YSCPL2M      200 bp    DNA             PLN       10-DEC-1984
DEFINITION  Yeast (S.cerevisiae) 2 micron plasmid (A-form) inverted repeat
            region.
ACCESSION   K01710
KEYWORDS    plasmid.
SOURCE      Yeast (Saccharomyces cerevisiae) 2 micron plasmid DNA.
```



```
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Endomycetales;
            Saccharomycetaceae.
REFERENCE   1  (bases 1 to 200)
  AUTHORS   Fagrelius,T.J. and Livingston,D.M.
  TITLE     Location of DNAase I sensitive cleavage sites in the yeast 2 mu-m
            plasmid DNA chromosome
  JOURNAL   J. Mol. Biol. 173, 1-13 (1984)
  MEDLINE   84138647
COMMENT     [1] examines whether cleavage sites are specific when the
            DNA-associated protein is stripped away and draws the conclusion
            that the specificity of DNAase I is dependent on the presence of
            nucleoprotein.

            NCBI gi: 172188
FEATURES             Location/Qualifiers
     source          1..200
                     /organism="Saccharomyces cerevisiae"
BASE COUNT       57 a     47 c     46 g     50 t
ORIGIN      103 bp upstream of XbaI site.

  Query Match          84.6%;  Score 22;  DB 43;  Length 200;
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db     90 gaagttcctattctctagaaagtataggaacttc 123
          |||||||||||  |         | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   39
LOCUS       PRS424       5616 bp    DNA   circular  SYN       24-MAY-1995
DEFINITION  Yeast episomal vector pRS424 with TRP1 marker, complete sequence.
ACCESSION   U03453
KEYWORDS    .
SOURCE      Cloning vector pRS424.
  ORGANISM  Cloning vector pRS424
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 5616)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 5616)
  AUTHORS   Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and
            Hieter,P.
  TITLE     Multifunctional yeast high-copy-number shuttle vectors
  JOURNAL   Gene 110 (1), 119-122 (1992)
  MEDLINE   92184105
REFERENCE   3  (bases 1 to 5616)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416324
FEATURES             Location/Qualifiers
     source          1..5616
                     /organism="cloning vector pRS424"
BASE COUNT     1513 a   1221 c   1356 g   1526 t
ORIGIN

  Query Match          84.6%;  Score 22;  DB 61;  Length 5616;
```

```
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   5169 gaagttcctatactttctagagaataggaacttc 5202
          |||||||||||  |        | ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34

RESULT   40
LOCUS       PRS426       5726 bp    DNA   circular  SYN       24-MAY-1995
DEFINITION  Yeast episomal vector pRS426 with URA3 marker, complete sequence.
ACCESSION   U03451
KEYWORDS    .
SOURCE      Cloning vector pRS426.
  ORGANISM  Cloning vector pRS426
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 5726)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 5726)
  AUTHORS   Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and
            Hieter,P.
  TITLE     Multifunctional yeast high-copy-number shuttle vectors
  JOURNAL   Gene 110 (1), 119-122 (1992)
  MEDLINE   92184105
REFERENCE   3  (bases 1 to 5726)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132 USA
COMMENT     NCBI gi: 416322
FEATURES             Location/Qualifiers
     source          1..5726
                     /organism="cloning vector pRS426"
BASE COUNT     1568 a   1246 c   1370 g   1542 t
ORIGIN

  Query Match           84.6%;  Score 22;  DB 61;  Length 5726;
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   5279 gaagttcctatactttctagagaataggaacttc 5312
          |||||||||||  |        | ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   41
LOCUS       CVU37458     6624 bp    DNA   circular  SYN       08-NOV-1995
DEFINITION  Yeast CUP1 expression-multicopy (2micron) cloning vector YRTAG300
            with the hemagglutinin tag sequence, complete sequence. selection).
ACCESSION   U37458
KEYWORDS    .
SOURCE      Cloning vector YRTAG300.
  ORGANISM  Cloning vector YRTAG300
            artificial sequence; cloning vectors.
REFERENCE   1  (bases 1 to 6624)
  AUTHORS   Lieberman,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Benjamin Lieberman, Pharmacology, Duke
            University, Research Drive, P.O. 3813, Durham, NC 27710, USA
COMMENT     NCBI gi: 1052969
```



```
FEATURES             Location/Qualifiers
     source          1..6624
                     /organism="Saccharomyces cerevisiae"
                     /note="based on pRS424 (TRP selection); includes CUP1
                     promoter, Cyc terminator and the hemagglutinin coding
                     region fused in frame to a start codon and and two
                     restriction sites (SstI and XhoI)"
     CDS             574..1434
                     /note="NCBI gi: 1052970"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     misc_feature    3040..3098
                     /note="HA TAG with start codon (EcoRI-HATAG-SSTI-XHOI)"
     CDS             complement(4514..5215)
                     /gene="TRP1"
                     /note="NCBI gi: 1052971"
                     /codon_start=1
                     /transl_table=11
                     /product="N-(5'-phosphoribosyl)-anthranilate isomerase"
                     /translation="MKHTKAAWSMSVINFTGSSGPLVKVCGLQSTEAAECALDSDADL
                     LGIICVPNRKRTIDPVIARKISSLVKAYKNSSGTPKYLVGVFRNQPKEDVLALVNDYG
                     IDIVQLHGDESWQEYQEFLGLPVIKRLVFPKDCNILLSAASQKPHSFIPLFDSEAGGT
                     GELLDWNSISDWVGRQESPESLHFMLAGGLTPENVGDALRLNGVIGVDVSGGVETNGV
                     KDSNKIANFVKNAKK"
BASE COUNT     1845 a   1543 c   1397 g   1839 t
ORIGIN

  Query Match           84.6%;  Score 22;  DB 84;  Length 6624;
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   6070 gaagttcctattctctagaaagtataggaacttc 6103
          |||||||||||  |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   42
ID   CV37458    standard; circular DNA; FUN; 6624 BP.
AC   U37458;
DT   10-NOV-1995 (Rel. 45, Created)
DT   10-NOV-1995 (Rel. 45, Last updated, Version 1)
DE   Yeast CUP1 expression-multicopy (2micron) cloning vector YRTAG300
DE   with the hemagglutinin tag sequence, complete sequence. selection).
KW   .
OS   Saccharomyces cerevisiae (yeast)
OC   Eukaryota; Plantae; Thallobionta; Eumycota; Hemiascomycetes;
OC   Endomycetales; Saccharomycetaceae.
RN   [1]
RP   1-6624
RA   Lieberman B.;
RT   ;
RL   Submitted (03-OCT-1995) to the EMBL/GenBank/DDBJ databases.
RL   Benjamin Lieberman, Pharmacology, Duke University, Research Drive,
RL   P.O. 3813, Durham, NC 27710, USA
CC   NCBI gi: 1052969
FH   Key             Location/Qualifiers
FH
```

```
FT   source          1..6624
FT                   /organism="Saccharomyces cerevisiae"
FT                   /note="based on pRS424 (TRP selection); includes CUP1
FT                   promoter, Cyc terminator and the hemagglutinin coding
FT                   region fused in frame to a start codon and and two
FT                   restriction sites (SstI and XhoI)"
FT   CDS             574..1434
FT                   /note="NCBI gi: 1052970"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /product="beta-lactamase"
FT                   /note="pid:g1052970"
FT   misc_feature    3040..3098
FT                   /note="HA TAG with start codon (EcoRI-HATAG-SSTI-XHOI)"
FT   CDS             complement(4514..5215)
FT                   /gene="TRP1"
FT                   /note="NCBI gi: 1052971"
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /product="N-(5'-phosphoribosyl)-anthranilate isomerase"
FT                   /note="pid:g1052971"
SQ   Sequence 6624 BP; 1845 A; 1543 C; 1397 G; 1839 T; 0 other;

  Query Match           84.6%;  Score 22;  DB 2;  Length 6624;
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   6070 gaagttcctattctctagaaagtataggaacttc 6103
          |||||||||||  |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   43
ID   CV33753    standard; circular DNA; SYN; 7834 BP.
AC   U33753;
DT   12-OCT-1995 (Rel. 45, Created)
DT   12-OCT-1995 (Rel. 45, Last updated, Version 1)
DE   Yeast episomal cloning vector pADNS, with ADH1 promoter, complete
DE   sequence.
KW   .
OS   Cloning vector pADNS
OC   Artificial sequences; Cloning vectors.
RN   [1]
RP   1-7834
RX   MEDLINE; 89264471.
RA   Colicelli J., Birchmeier C., Michaeli T., O'Neill K., Riggs M.,
RA   Wigler M.;
RT   "Isolation and characterization of a mammalian gene encoding a
RT   high-affinity cAMP phosphodiesterase";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:3599-3603(1989).
RN   [2]
RP   1-7834
RA   Stillman D.J.;
RT   ;
RL   Submitted (10-AUG-1995) to the EMBL/GenBank/DDBJ databases.
RL   David J. Stillman, Division of Molecular Biology and Genetics,
RL   Department of Oncological Sciences, University of Utah Heath
RL   Sciences Center, Salt Lake City, UT 84132, USA
CC   NCBI gi: 988314
FH   Key             Location/Qualifiers
FH
FT   source          1..7834
FT                   /organism="Cloning vector pADNS"
FT   misc_feature    1..38
```



```
FT                   /note="polylinker"
FT   terminator      39..505
FT                   /gene="ADH1"
FT   promoter        6381..7834
FT                   /gene="ADH1"
SQ   Sequence 7834 BP; 2192 A; 1672 C; 1646 G; 2324 T; 0 other;

  Query Match           84.6%;  Score 22;  DB 15;  Length 7834;
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   5763 gaagttcctatactttctagagaataggaacttc 5796
          |||||||||||  |        | ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   44
ID   A19996     standard; DNA; SYN; 7859 BP.
AC   A19996;
DT   14-JUL-1995 (Rel. 44, Created)
DT   14-JUL-1995 (Rel. 44, Last updated, Version 1)
DE   SEQ ID NO: 4; Synthetic plasmid pSW6.
KW   .
OS   None
OC   Artificial sequences.
RN   [1]
RA   ;
RT   "PROTEINS AND NUCLEIC ACIDS";
RL   Patent number WO9109125-A/4, 27-JUN-1991.
FH   Key             Location/Qualifiers
FH
FT   source          1..7859
FT                   /organism="Artificial sequences"
SQ   Sequence 7859 BP; 2317 A; 1656 C; 1600 G; 2286 T; 0 other;

  Query Match           84.6%;  Score 22;  DB 88;  Length 7859;
  Best Local Similarity  70.6%;  Pred. No. 1.34e-04;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          |||||||||||  |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   45
LOCUS       A18079       7984 bp    DNA             PAT       22-APR-1994
DEFINITION  yeast expression vector pSW6 seq ID No: 19.
ACCESSION   A18079
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 7984)
  AUTHORS   .
  TITLE     PHARMACEUTICALLY ACTIVE PROTEINS COMPRISING AN ACTIVE PROTEIN AND
            AN INTEGRINAFFINITY SEQUENCE
  JOURNAL   Patent: WO 9207874-A 33 14-MAY-1992;

COMMENT     NCBI gi: 513171
FEATURES             Location/Qualifiers
     source          1..7984
                     /organism="Artificial sequences"
                     /lab_host="yeast expression vector pSW6"
BASE COUNT     2345 a   1695 c   1638 g   2306 t
```

ORIGIN

Query Match 84.6%; Score 22; DB 34; Length 7984;
Best Local Similarity 70.6%; Pred. No. 1.34e-04;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          |||||||||||  |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1
```

Search completed: Tue May 14 13:58:25 1996
Job time : 534 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue May 14 13:58:43 1996; MasPar time 4.93 Seconds
458.714 Million cell updates/sec

Tabular output not generated.

Title: >FLP
Description: (1-34) from frt.seq
Perfect Score: 26
N.A. Sequence: 1 gaagttcctattcNNNNNNNNgtataggaacttc 34
Comp: cttcaaggataagNNNNNNNNcatatccttgaag

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 84802 seqs, 33246950 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16

Statistics: Mean 5.329; Variance 3.136; scale 1.699

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
c	1	26	100.0	54	12	Q67140	Complete FRT site lac	5.30e-06
	2	26	100.0	1340	15	Q93078	Neomycin-resistance c	5.30e-06
	3	26	100.0	7859	7	Q44265	pSW6 for expression o	5.30e-06
	4	26	100.0	7859	2	Q12154	Shuttle vector pSW6.	5.30e-06
	5	26	100.0	7984	4	Q25185	pSW6 expression vecto	5.30e-06
	6	25	96.2	33	5	Q29100	Sequence of FLP recom	2.13e-05
c	7	24	92.3	41	12	Q67141	Partial FRT site lack	8.46e-05
	8	22	84.6	54	12	Q67140	Complete FRT site lac	1.28e-03
c	9	22	84.6	91	9	Q51746	Oligonucleotide probe	1.28e-03

c	10	22	84.6	1340	15	Q93078	Neomycin-resistance c	1.28e-03
c	11	22	84.6	7859	2	Q12154	Shuttle vector pSW6.	1.28e-03
c	12	22	84.6	7859	7	Q44265	pSW6 for expression o	1.28e-03
c	13	22	84.6	7984	4	Q25185	pSW6 expression vecto	1.28e-03
c	14	21	80.8	33	5	Q29100	Sequence of FLP recom	4.88e-03
	15	20	76.9	41	12	Q67141	Partial FRT site lack	1.83e-02
	16	20	76.9	91	9	Q51746	Oligonucleotide probe	1.83e-02
c	17	16	61.5	204	1	N81164	Base substituted E.co	2.91e+00
	18	16	61.5	4093	9	Q49264	ced-4.	2.91e+00
	19	15	57.7	42	12	Q67134	DNA primer used for c	9.70e+00
c	20	15	57.7	1047	2	Q10572	Human Natriuretic Pep	9.70e+00
c	21	15	57.7	1971	10	Q56791	cDNA encoding recepto	9.70e+00
c	22	15	57.7	3249	12	Q71367	E.coli/S.cerevisiae s	9.70e+00
c	23	15	57.7	3400	12	Q71366	E.coli/S.cerevisiae s	9.70e+00
c	24	15	57.7	5211	13	Q77789	Pre-pro-cobra C3 codi	9.70e+00
	25	15	57.7	6824	7	Q39050	K.lactis/S. cerevisae	9.70e+00
	26	14	53.8	204	1	N81164	Base substituted E.co	3.13e+01
c	27	14	53.8	498	3	N50034	Sequence encoding new	3.13e+01
	28	14	53.8	501	3	N50025	Sequence encoding new	3.13e+01
	29	14	53.8	501	3	N50023	Sequence encoding new	3.13e+01
	30	14	53.8	501	3	N50026	Sequence encoding new	3.13e+01
	31	14	53.8	501	3	N50031	Sequence encoding new	3.13e+01
	32	14	53.8	501	3	N50029	Sequence encoding new	3.13e+01
	33	14	53.8	501	3	N50027	Sequence encoding new	3.13e+01
	34	14	53.8	501	3	N50032	Sequence encoding new	3.13e+01
	35	14	53.8	501	3	N50024	Sequence encoding new	3.13e+01
	36	14	53.8	501	3	N50033	Sequence encoding new	3.13e+01
	37	14	53.8	1561	4	Q25420	Encodes human liver c	3.13e+01
c	38	14	53.8	1997	8	Q47839	Human interleukin 9 r	3.13e+01
	39	14	53.8	3249	12	Q71367	E.coli/S.cerevisiae s	3.13e+01
	40	14	53.8	3400	12	Q71366	E.coli/S.cerevisiae s	3.13e+01
c	41	14	53.8	4093	9	Q49264	ced-4.	3.13e+01
c	42	14	53.8	10097	3	Q24802	SIVmac239 nef-deletio	3.13e+01
c	43	14	53.8	10279	3	Q22487	SIVmac239 proviral ge	3.13e+01
	44	14	53.8	12151	10	Q54676	Rice starch branching	3.13e+01
	45	14	53.8	12151	12	Q62137	Rice starch branching	3.13e+01

ALIGNMENTS

RESULT 1
ID Q67140 standard; DNA; 54 BP.
AC Q67140;
DT 22-MAR-1995 (first entry)
DE Complete FRT site lacking additional 5 FLP binding sites.
KW Maize; Zea mays; cereal; grass; protoplast; FLP; ss.
OS Synthetic.
PN WO9417176-A.
PD 04-AUG-1994.
PF 27-JAN-1994; U00927.
PR 29-JAN-1993; US-010997.
PA (PURD) PURDUE RES FOUND.
PI Hodges TK, Lyznik LA;
DR WPI; 94-264090/32.
PT DNA constructs - for creating transgenic eukaryotic cells
PS Disclosure; Page 51 79pp; English.
CC This sequence is of the complete FRT site which is ligated into the
CC BglIII site of the ubiqutin first exon. This FRP site lacks
CC additional 5 FLP protein binding sites, and has application in the
CC construction of transgenic eukaryotic cells.
SQ Sequence 54 BP; 18 A; 9 C; 11 G; 16 T;

Query Match 100.0%; Score 26; DB 12; Length 54;
Best Local Similarity 76.5%; Pred. No. 5.30e-06;

```
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db      6 gaagttcctatactttctagagaataggaacttc 39
          ||||||||||||        ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT    2
ID   Q93078 standard; cDNA; 1340 BP.
AC   Q93078;
DT   10-DEC-1995  (first entry)
DE   Neomycin-resistance cassette.
KW   Alpha-1,3-galactosyltransferase; alpha-1,3-GalT; transgenic animal;
KW   mouse; hyperacute rejection; xerotransplantation; donor organ;
KW   allograft rejection; Gal epitope; gene disruption;
KW   homologous recombination; knock-out; neomycin-resistance; ss.
OS   Not specified.
FH   Key             Location/Qualifiers
FT   misc_feature    1..28
FT   /*tag=  a
FT   /function= linker sequence
FT   misc_feature    29..104
FT   /*tag=  b
FT   /function= FLP recombinase target site
FT   enhancer        105..249
FT   /*tag=  c
FT   /function= polyoma virus enhancer repeats
FT   promoter        250..385
FT   /*tag=  d
FT   /function= herpes simplex virus tyrosine-kinase
FT   promoter
FT   CDS             385..1188
FT   /*tag=  e
FT   /product=  neomycin-phosphotransferase
FT   polyA_signal    1189..1249
FT   /*tag=  f
FT   /function= herpes simplex virus tyrosine-kinase
FT   polyA signal
FT   misc_feature    1250..1310
FT   /*tag=  g
FT   /function= FLP recombinase target site
FT   misc_feature    1311..1340
FT   /*tag=  h
FT   /function= linker sequences
PN   WO9520661-A1.
PD   03-AUG-1995.
PF   27-JAN-1995; IB0088.
PR   27-JAN-1994; US-188607.
PR   26-JAN-1995; US-188607.
PA   (BRES-) BRESATEC LTD.
PA   (SVIN-) ST VINCENT'S HOSPITAL MELBOURNE LTD.
PI   Crawford RJ,  Dapice AJF,  Pearse MJ,  Rathjen PD;
PI   Robbins AJ;
DR   WPI; 95-275446/36.
PT   New alpha-1,3-galactosyltransferase and leukaemia inhibitor factor
PT   - corresp. DNA and nucleic acid constructs for inactivating the
PT   transferase gene; for eliminating hyperacute region in human
PT   transplants
PS   Disclosure; Fig.16a-16b; 184pp; English.
CC   The neomycin-resistance cassette given in Q93078 was used in the
CC   development of a DNA construct (pNeo-alpha-GT10.8B) used to
CC   interrupt the mouse alpha-1,3-GalT gene by means of homologous
CC   recombination as a means of suppressing the GAL epitope.
SQ   Sequence   1340 BP;    285 A;    362 C;    391 G;    302 T;
```



```
  Query Match          100.0%;  Score 26;  DB 15;  Length 1340;
  Best Local Similarity  76.5%;  Pred. No. 5.30e-06;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     48 gaagttcctattctctagaaagtataggaacttc 81
          ||||||||||||        ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    3
ID   Q44265 standard; DNA; 7859 BP.
AC   Q44265;
DT   23-NOV-1993  (first entry)
DE   pSW6 for expression of LD78 synthetic gene.
KW   SCI; stem cell inhibition; LD78; ACT2; MIP-1alpha;
KW   macrophage inflammatory protein; multimer; tumour therapy;
KW   psoriasis; hyperproliferation; yeast expression vector;
KW   circular; ds.
OS   Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   misc_difference 1773
FT   /*tag=  a
FT   /note= "base illegible in the specification"
PN   WO9313206-A.
PD   08-JUL-1993.
PF   23-DEC-1992; G02390.
PR   23-DEC-1991; GB-027319.
PR   14-OCT-1992; GB-021587.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI   Craig S,  Czaplewski LG,  Edwards RM,  Gilbert RJ;
PI   Hunter MG;
DR   WPI; 93-227322/28.
PT   Protein with stem cell inhibition activity, e.g. LD78 or MIP-1
PT   alpha - unable to form stable multimer higher than dodecamer,
PT   providing better tissue penetration
PS   Disclosure; Page 159-168; 294pp; English.
CC   An expression vector was designed to enable secretion of LD78 to
CC   the extracellular medium after expression in S. cerevisiae.
CC   Secretion aids purification and rapid analysis of LD78.
CC   The secretion signals from the yeast mating type factor alpha were
CC   used to direct export of the LD78 protein.  The yeast expression
CC   vector pSW6 (NCIMB 40326) is based on the 2 micron circle from
CC   S. cerevisiae.
SQ   Sequence   7859 BP;    2317 A;    1667 C;    1585 G;    2289 T;

  Query Match          100.0%;  Score 26;  DB 7;  Length 7859;
  Best Local Similarity  76.5%;  Pred. No. 5.30e-06;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          ||||||||||||        ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    4
ID   Q12154 standard; DNA; 7859 BP.
AC   Q12154;
DT   17-SEP-1991  (first entry)
DE   Shuttle vector pSW6.
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; streptokinase; plasmid; circular; ss.
OS   Synthetic.
PN   WO9109125-A.
```

```
PD   27-JUN-1991.
PF   07-DEC-1990; G01911.
PR   07-DEC-1989; GB-027722.
PR   07-DEC-1990; WO-G01911.
PA   (BRBI-) BRIT BIO-TECHN LTD.
PI   Dawson KM, Hunter MG, Czapleswski LG;
DR   WPI; 91-208151/28.
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
PS   Disclosure; Page 71; 115pp; English.
CC   The vector is based on the 2u circle from S. cerevisiae.  It is
CC   deposited in S. cerevisiae strain BJ2168 as NCIMB 40326.  It is a
CC   shuttle vector capable of replcation in both E. coli and S. cere-
CC   visiae and contains origins of replication for both, the leu2 gene
CC   (selectable marker), and an ampicillin resistant locus. The E. coli
CC   sequences are derived from E. coli ColE1-based replicon pAT153. The
CC   vector contains an alpha factor pre-pro-peptide gene fused in frame
CC   to the gene for epidermal growth factor (EGF).  The expression of
CC   this fusion is under control of a galactodse regulated promoter
CC   which contains hybrid DNA from S. cerevisiae GAL 1-10 promoter and
CC   the S. cerevisiae phosphoglycerate kinase (PGK) promoter. The EGF
CC   gene can be excised by digestion with HindIII and BamHI. The plas-
CC   mid was used for the expression of a synthetic hirudin HV-1 gene
CC   in E. coli K12 HW87.  The plasmid can be used to construct ex-
CC   pression vectors in which the hirudin gene is linked to a second
CC   gene encoding e.g. another hirudin protein, streptokinase or a
CC   streptokinase-like protein, via a linking peptide.  This peptide
CC   link contains a cleavage site for e.g. factor X or thrombin which
CC   can be cleaved, releasing the individual proteins which have anti-
CC   thrombotic activity.  The enzymes which cleave the fusion protein
CC   are present at the site of the target thrombus so the active agents
CC   are released specifically at the place where clot formation is
CC   occurring.
CC   See also Q12153-Q12156, Q12158-Q12162 and Q12490.
SQ   Sequence   7859 BP;    2317 A;    1656 C;    1600 G;    2286 T;

  Query Match          100.0%;  Score 26;  DB 2;  Length 7859;
  Best Local Similarity  76.5%;  Pred. No. 5.30e-06;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    5
ID   Q25185 standard; DNA; 7984 BP.
AC   Q25185;
DT   18-NOV-1992  (first entry)
DE   pSW6 expression vector.
KW   Escherichia coli; 2 micron circle; shuttle vector; leu2; EGF;
KW   ampicilin resistant locus; epidermal growth factor; GAL 1-10;
KW   phosphoglycerate kinase promoter; PGK; BamHI; HindIII; ss.
OS   Saccharomyces cerevisae.
PN   WO9207874-A.
PD   14-MAY-1992.
PF   23-OCT-1991; G01860.
PR   24-OCT-1990; GB-023149.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI   Dawson KM, Edwards RM, Fallon A;
DR   WPI; 92-183627/22.
PT   New proteins comprising active protein and integrin-affinity
PT   sequence - are antithrombotics useful in treating and preventing
```



```
PT   myocardial infarction, stroke, pulmonary embolism and deep vein
PT   thrombosis
PS   Disclosure; Page 67; 101pp; English.
CC   The sequence given is the yeast expression vector pSW6.  It is based
CC   on the 2 micron circle from Saccharomyces cerevisae.  It is a shuttle
CC   vector capable of replication in both S. cerevisae and Escherichia
CC   coli as it contains the origin of replication for both organisms.  It
CC   also contains the leu2 gene (a yeast selectable marker) and the
CC   amipicillin resistant locus for selection of plasmid maintenance in E.
CC   coli.  This vector has enhanced ability for passage through E.coli and
CC   this greatly facilitates genetic manipulation with this vector.  pSW6
CC   contains contains an alpha- factor pre-pro peptide fused in-frame to
CC   epidermal growth factor (EGF).  The expression of this fusion is under
CC   the control of an efficient galactose regulated promoter which contains
CC   hybrid DNA sequences from the S. cerevisae GAL 1-10 promoter and the S.
CC   cerevisae phosphoglycerate kinase (PGK) promoter.  Transcription is
CC   terminated in this vector by the natural yeast PGK terminator.  The EGF
CC   gene in pSW6 can be removed by digestion with HindIII and BamHI.  This
CC   removes  DNA encoding both EGF and 5 amino acids from the C-terminus of
CC   the alpha-factor pro-peptide.  Genes to be inserted into the pSW6
CC   expression vector must therefore have the general composition: HindIII
CC   site-alpha-factor adapter-gene-BamHI site.
SQ   Sequence   7984 BP;    2348 A;    1698 C;    1635 G;    2303 T;

  Query Match          100.0%;  Score 26;  DB 4;  Length 7984;
  Best Local Similarity  76.5%;  Pred. No. 5.30e-06;
  Matches    26;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          ||||||||||||||        ||||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    6
ID   Q29100 standard; DNA; 33 BP.
AC   Q29100;
DT   25-FEB-1992  (first entry)
DE   Sequence of FLP recombination target site
KW   FLP recombinase; site-specific integration system; gene activation;
KW   gene inactivation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    14..21
FT   /*tag=  a
FT   /label= spacer
PN   WO9215694-A.
PD   17-SEP-1992.
PF   06-MAR-1992; U01899.
PR   08-MAR-1991; US-666252.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ogorman SV, Wahl GM;
DR   WPI; 92-331739/40.
PT   FLP-mediated gene modification in mammalian cells - giving
PT   precise modification by recombination and can be used to alter
PT   transgenes for therapeutic purposes and analysis of development
PS   Claim 33; Page 40; 49pp; English.
CC   FLP recombinase is a protein which catalyses a site-specific
CC   recombination reaction that is involved in amplifying the copy
CC   number of the 2-mu plasmid of S. cerevisiae during DNA replication.
CC   The inventors claim a mammalian recombination system in which the
CC   FLP recombinase is pref. Q29101. The FLP recombination target site
CC   (FRT) has been identified as minimally comprising two 13 base-pair
CC   repeats, separated by an 8 base-pair spacer (see Q29100). The
CC   nucleotides in the spacer region can be replaced with any other
```

```
CC   combination of nucleotides so long as the two 13 base-pair repeats
CC   are separated by 8 nucleotides. NB, in the claims the sequence of
CC   the FRT has only 12 base pairs on the 3' end of the spacer. The
CC   apparently missing base would be C.
SQ   Sequence   33 BP;    11 A;    5 C;    6 G;    11 T;

  Query Match           96.2%;  Score 25;  DB 5;  Length 33;
  Best Local Similarity  75.8%;  Pred. No. 2.13e-05;
  Matches    25;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db      1 gaagttcctattctctagaaagtataggaacttt 33
          |||||||||||||        ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttt 33


RESULT    7
ID   Q67141 standard; DNA; 41 BP.
AC   Q67141;
DT   22-MAR-1995  (first entry)
DE   Partial FRT site lacking additional 5 FLP binding sites.
KW   Maize; Zea mays; cereal; grass; protoplast; FLP; ss.
OS   Synthetic.
PN   WO9417176-A.
PD   04-AUG-1994.
PF   27-JAN-1994; U00927.
PR   29-JAN-1993; US-010997.
PA   (PURD ) PURDUE RES FOUND.
PI   Hodges TK,  Lyznik LA;
DR   WPI; 94-264090/32.
PT   DNA constructs - for creating transgenic eukaryotic cells
PS   Disclosure; Page 51 79pp; English.
CC   This sequence is of the partial FRT site which is ligated into the
CC   BglIII site of the ubiqutin first exon.  This FRP site lacks
CC   additional 5 FLP protein binding sites, and has application in the
CC   construction of transgenic eukaryotic cells.
SQ   Sequence   41 BP;    13 A;    7 C;    8 G;    13 T;

  Query Match           92.3%;  Score 24;  DB 12;  Length 41;
  Best Local Similarity  75.0%;  Pred. No. 8.46e-05;
  Matches    24;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db      5 agttcctatactttctagagaataggaacttc 36
          ||||||||||||        ||||||||||||||
Cp     32 agttcctatacNNNNNNNNgaataggaacttc 1


RESULT    8
ID   Q67140 standard; DNA; 54 BP.
AC   Q67140;
DT   22-MAR-1995  (first entry)
DE   Complete FRT site lacking additional 5 FLP binding sites.
KW   Maize; Zea mays; cereal; grass; protoplast; FLP; ss.
OS   Synthetic.
PN   WO9417176-A.
PD   04-AUG-1994.
PF   27-JAN-1994; U00927.
PR   29-JAN-1993; US-010997.
PA   (PURD ) PURDUE RES FOUND.
PI   Hodges TK,  Lyznik LA;
DR   WPI; 94-264090/32.
PT   DNA constructs - for creating transgenic eukaryotic cells
PS   Disclosure; Page 51 79pp; English.
CC   This sequence is of the complete FRT site which is ligated into the
CC   BglIII site of the ubiqutin first exon.  This FRP site lacks
```



```
CC   additional 5 FLP protein binding sites, and has application in the
CC   construction of transgenic eukaryotic cells.
SQ   Sequence   54 BP;    18 A;    9 C;    11 G;    16 T;

  Query Match           84.6%;  Score 22;  DB 12;  Length 54;
  Best Local Similarity  70.6%;  Pred. No. 1.28e-03;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db      6 gaagttcctatactttctagagaataggaacttc 39
          |||||||||||  |        | ||||||||||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT    9
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match           84.6%;  Score 22;  DB 9;  Length 91;
  Best Local Similarity   0.0%;  Pred. No. 1.28e-03;
  Matches     0;  Conservative   24;  Mismatches  10;  Indels   0;  Gaps   0;

Db     19 vhvvshhhsvhhvvhhvhvsvvvvhhvvhvvhhv 52
          :::: ::: ::::        ::::::::::::::
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   10
ID   Q93078 standard; cDNA; 1340 BP.
AC   Q93078;
DT   10-DEC-1995  (first entry)
DE   Neomycin-resistance cassette.
KW   Alpha-1,3-galactosyltransferase; alpha-1,3-GalT; transgenic animal;
KW   mouse; hyperacute rejection; xerotransplantation; donor organ;
KW   allograft rejection; Gal epitope; gene disruption;
KW   homologous recombination; knock-out; neomycin-resistance; ss.
OS   Not specified.
FH   Key             Location/Qualifiers
FT   misc_feature    1..28
FT   /*tag=  a
FT   /function= linker sequence
FT   misc_feature    29..104
FT   /*tag=  b
```

```
FT   /function= FLP recombinase target site
FT   enhancer        105..249
FT   /*tag=  c
FT   /function= polyoma virus enhancer repeats
FT   promoter        250..385
FT   /*tag=  d
FT   /function= herpes simplex virus tyrosine-kinase
FT   promoter
FT   CDS             385..1188
FT   /*tag=  e
FT   /product=  neomycin-phosphotransferase
FT   polyA_signal    1189..1249
FT   /*tag=  f
FT   /function= herpes simplex virus tyrosine-kinase
FT   polyA signal
FT   misc_feature    1250..1310
FT   /*tag=  g
FT   /function= FLP recombinase target site
FT   misc_feature    1311..1340
FT   /*tag=  h
FT   /function= linker sequences
PN   WO9520661-A1.
PD   03-AUG-1995.
PF   27-JAN-1995; IB0088.
PR   27-JAN-1994; US-188607.
PR   26-JAN-1995; US-188607.
PA   (BRES-) BRESATEC LTD.
PA   (SVIN-) ST VINCENT'S HOSPITAL MELBOURNE LTD.
PI   Crawford RJ,  Dapice AJF,  Pearse MJ,  Rathjen PD;
PI   Robbins AJ;
DR   WPI; 95-275446/36.
PT   New alpha-1,3-galactosyltransferase and leukaemia inhibitor factor
PT   - corresp. DNA and nucleic acid constructs for inactivating the
PT   transferase gene; for eliminating hyperacute region in human
PT   transplants
PS   Disclosure; Fig.16a-16b; 184pp; English.
CC   The neomycin-resistance cassette given in Q93078 was used in the
CC   development of a DNA construct (pNeo-alpha-GT10.8B) used to
CC   interrupt the mouse alpha-1,3-GalT gene by means of homologous
CC   recombination as a means of suppressing the GAL epitope.
SQ   Sequence   1340 BP;    285 A;    362 C;    391 G;    302 T;

  Query Match           84.6%;  Score 22;  DB 15;  Length 1340;
  Best Local Similarity  70.6%;  Pred. No. 1.28e-03;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db     48 gaagttcctattctctagaaagtataggaacttc 81
          |||||||||||| |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   11
ID   Q12154 standard; DNA; 7859 BP.
AC   Q12154;
DT   17-SEP-1991  (first entry)
DE   Shuttle vector pSW6.
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; streptokinase; plasmid; circular; ss.
OS   Synthetic.
PN   WO9109125-A.
PD   27-JUN-1991.
PF   07-DEC-1990; G01911.
PR   07-DEC-1989; GB-027722.
PR   07-DEC-1990; WO-G01911.
```



```
PA   (BRBI-) BRIT BIO-TECHN LTD.
PI   Dawson KM, Hunter MG, Czapleswski LG;
DR   WPI; 91-208151/28.
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
PS   Disclosure; Page 71; 115pp; English.
CC   The vector is based on the 2u circle from S. cerevisiae.  It is
CC   deposited in S. cerevisiae strain BJ2168 as NCIMB 40326.  It is a
CC   shuttle vector capable of replcation in both E. coli and S. cere-
CC   visiae and contains origins of replication for both, the leu2 gene
CC   (selectable marker), and an ampicillin resistant locus. The E. coli
CC   sequences are derived from E. coli ColE1-based replicon pAT153. The
CC   vector contains an alpha factor pre-pro-peptide gene fused in frame
CC   to the gene for epidermal growth factor (EGF).  The expression of
CC   this fusion is under control of a galactodse regulated promoter
CC   which contains hybrid DNA from S. cerevisiae GAL 1-10 promoter and
CC   the S. cerevisiae phosphoglycerate kinase (PGK) promoter. The EGF
CC   gene can be excised by digestion with HindIII and BamHI. The plas-
CC   mid was used for the expression of a synthetic hirudin HV-1 gene
CC   in E. coli K12 HW87.  The plasmid can be used to construct ex-
CC   pression vectors in which the hirudin gene is linked to a second
CC   gene encoding e.g. another hirudin protein, streptokinase or a
CC   streptokinase-like protein, via a linking peptide.  This peptide
CC   link contains a cleavage site for e.g. factor X or thrombin which
CC   can be cleaved, releasing the individual proteins which have anti-
CC   thrombotic activity.  The enzymes which cleave the fusion protein
CC   are present at the site of the target thrombus so the active agents
CC   are released specifically at the place where clot formation is
CC   occurring.
CC   See also Q12153-Q12156, Q12158-Q12162 and Q12490.
SQ   Sequence   7859 BP;    2317 A;    1656 C;    1600 G;    2286 T;

  Query Match           84.6%;  Score 22;  DB 2;  Length 7859;
  Best Local Similarity  70.6%;  Pred. No. 1.28e-03;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          |||||||||||| |        | ||||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   12
ID   Q44265 standard; DNA; 7859 BP.
AC   Q44265;
DT   23-NOV-1993  (first entry)
DE   pSW6 for expression of LD78 synthetic gene.
KW   SCI; stem cell inhibition; LD78; ACT2; MIP-1alpha;
KW   macrophage inflammatory protein; multimer; tumour therapy;
KW   psoriasis; hyperproliferation; yeast expression vector;
KW   circular; ds.
OS   Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   misc_difference 1773
FT   /*tag=  a
FT   /note= "base illegible in the specification"
PN   WO9313206-A.
PD   08-JUL-1993.
PF   23-DEC-1992; G02390.
PR   23-DEC-1991; GB-027319.
PR   14-OCT-1992; GB-021587.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI   Craig S,  Czaplewski LG,  Edwards RM,  Gilbert RJ;
PI   Hunter MG;
```

```
DR   WPI; 93-227322/28.
PT   Protein with stem cell inhibition activity, e.g. LD78 or MlP-1
PT   alpha - unable to form stable multimer higher than dodecamer,
PT   providing better tissue penetration
PS   Disclosure; Page 159-168; 294pp; English.
CC   An expression vector was designed to enable secretion of LD78 to
CC   the extracellular medium after expression in S. cerevisiae.
CC   Secretion aids purification and rapid analysis of LD78.
CC   The secretion signals from the yeast mating type factor alpha were
CC   used to direct export of the LD78 protein.  The yeast expression
CC   vector pSW6 (NCIMB 40326) is based on the 2 micron circle from
CC   S. cerevisiae.
SQ   Sequence   7859 BP;    2317 A;    1667 C;    1585 G;    2289 T;

  Query Match           84.6%;  Score 22;  DB 7;  Length 7859;
  Best Local Similarity  70.6%;  Pred. No. 1.28e-03;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          |||||||||||  |        | ||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   13
ID   Q25185 standard; DNA; 7984 BP.
AC   Q25185;
DT   18-NOV-1992  (first entry)
DE   pSW6 expression vector.
KW   Escherichia coli; 2 micron circle; shuttle vector; leu2; EGF;
KW   ampicilin resistant locus; epidermal growth factor; GAL 1-10;
KW   phosphoglycerate kinase promoter; PGK; BamHI; HindIII; ss.
OS   Saccharomyces cerevisae.
PN   WO9207874-A.
PD   14-MAY-1992.
PF   23-OCT-1991; G01860.
PR   24-OCT-1990; GB-023149.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI   Dawson KM, Edwards RM, Fallon A;
DR   WPI; 92-183627/22.
PT   New proteins comprising active protein and integrin-affinity
PT   sequence - are antithrombotics useful in treating and preventing
PT   myocardial infarction, stroke, pulmonary embolism and deep vein
PT   thrombosis
PS   Disclosure; Page 67; 101pp; English.
CC   The sequence given is the yeast expression vector pSW6.  It is based
CC   on the 2 micron circle from Saccharomyces cerevisae.  It is a shuttle
CC   vector capable of replication in both S. cerevisae and Escherichia
CC   coli as it contains the origin of replication for both organisms.  It
CC   also contains the leu2 gene (a yeast selectable marker) and the
CC   amipicillin resistant locus for selection of plasmid maintenance in E.
CC   coli.  This vector has enhanced ability for passage through E.coli and
CC   this greatly facilitates genetic manipulation with this vector.  pSW6
CC   contains contains an alpha- factor pre-pro peptide fused in-frame to
CC   epidermal growth factor (EGF).  The expression of this fusion is under
CC   the control of an efficient galactose regulated promoter which contains
CC   hybrid DNA sequences from the S. cerevisae GAL 1-10 promoter and the S.
CC   cerevisae phosphoglycerate kinase (PGK) promoter.  Transcription is
CC   terminated in this vector by the natural yeast PGK terminator.  The EGF
CC   gene in pSW6 can be removed by digestion with HindIII and BamHI.  This
CC   removes  DNA encoding both EGF and 5 amino acids from the C-terminus of
CC   the alpha-factor pro-peptide.  Genes to be inserted into the pSW6
CC   expression vector must therefore have the general composition: HindIII
CC   site-alpha-factor adapter-gene-BamHI site.
SQ   Sequence   7984 BP;    2348 A;    1698 C;    1635 G;    2303 T;
```



```
  Query Match           84.6%;  Score 22;  DB 4;  Length 7984;
  Best Local Similarity  70.6%;  Pred. No. 1.28e-03;
  Matches    24;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   3131 gaagttcctattctctagaaagtataggaacttc 3164
          |||||||||||  |        | ||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   14
ID   Q29100 standard; DNA; 33 BP.
AC   Q29100;
DT   25-FEB-1992  (first entry)
DE   Sequence of FLP recombination target site
KW   FLP recombinase; site-specific integration system; gene activation;
KW   gene inactivation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    14..21
FT   /*tag=  a
FT   /label= spacer
PN   WO9215694-A.
PD   17-SEP-1992.
PF   06-MAR-1992; U01899.
PR   08-MAR-1991; US-666252.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ogorman SV, Wahl GM;
DR   WPI; 92-331739/40.
PT   FLP-mediated gene modification in mammalian cells - giving
PT   precise modification by recombination and can be used to alter
PT   transgenes for therapeutic purposes and analysis of development
PS   Claim 33; Page 40; 49pp; English.
CC   FLP recombinase is a protein which catalyses a site-specific
CC   recombination reaction that is involved in amplifying the copy
CC   number of the 2-mu plasmid of S. cerevisiae during DNA replication.
CC   The inventors claim a mammalian recombination system in which the
CC   FLP recombinase is pref. Q29101. The FLP recombination target site
CC   (FRT) has been identified as minimally comprising two 13 base-pair
CC   repeats, separated by an 8 base-pair spacer (see Q29100). The
CC   nucleotides in the spacer region can be replaced with any other
CC   combination of nucleotides so long as the two 13 base-pair repeats
CC   are separated by 8 nucleotides. NB, in the claims the sequence of
CC   the FRT has only 12 base pairs on the 3' end of the spacer. The
CC   apparently missing base would be C.
SQ   Sequence   33 BP;    11 A;    5 C;    6 G;    11 T;

  Query Match           80.8%;  Score 21;  DB 5;  Length 33;
  Best Local Similarity  69.7%;  Pred. No. 4.88e-03;
  Matches    23;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db      1 gaagttcctattctctagaaagtataggaactt 33
          |||||||||||  |        | ||||||||||
Cp     34 gaagttcctatacNNNNNNNNgaataggaactt 2


RESULT   15
ID   Q67141 standard; DNA; 41 BP.
AC   Q67141;
DT   22-MAR-1995  (first entry)
DE   Partial FRT site lacking additional 5 FLP binding sites.
KW   Maize; Zea mays; cereal; grass; protoplast; FLP; ss.
OS   Synthetic.
PN   WO9417176-A.
```

```
PD   04-AUG-1994.
PF   27-JAN-1994; U00927.
PR   29-JAN-1993; US-010997.
PA   (PURD ) PURDUE RES FOUND.
PI   Hodges TK,  Lyznik LA;
DR   WPI; 94-264090/32.
PT   DNA constructs - for creating transgenic eukaryotic cells
PS   Disclosure; Page 51 79pp; English.
CC   This sequence is of the partial FRT site which is ligated into the
CC   BglIII site of the ubiqutin first exon.  This FRP site lacks
CC   additional 5 FLP protein binding sites, and has application in the
CC   construction of transgenic eukaryotic cells.
SQ   Sequence   41 BP;    13 A;    7 C;    8 G;    13 T;

  Query Match           76.9%;  Score 20;  DB 12;  Length 41;
  Best Local Similarity  68.8%;  Pred. No. 1.83e-02;
  Matches    22;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db      5 agttcctatactttctagagaataggaacttc 36
          ||||||||| |        | ||||||||||
Qy      3 agttcctattcNNNNNNNNgtataggaacttc 34


RESULT   16
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match           76.9%;  Score 20;  DB 9;  Length 91;
  Best Local Similarity   0.0%;  Pred. No. 1.83e-02;
  Matches     0;  Conservative   23;  Mismatches  11;  Indels   0;  Gaps   0;

Db     12 svhsyyvvhvvshhhsvhhvvhhvhvsvvvvhhv 45
          ::::::::::  :         ::::::::::::
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   17
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990  (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
```



```
OS   Escherichia coli.
FH   Key             Location/Qualifiers
FT   misc_feature    19..69
FT   /*tag=  a
FT   /function=multiple cloning site
FT   primer_bind     187..204
FT   /*tag=  b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI; 88-279927/40.
PT   Introducing random point mutations into nucleic acods -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; p; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence   204 BP;   21 A;   47 C;   17 G;   11 T;   108 Others;

  Query Match           61.5%;  Score 16;  DB 1;  Length 204;
  Best Local Similarity  18.5%;  Pred. No. 2.91e+00;
  Matches     5;  Conservative   15;  Mismatches   7;  Indels   0;  Gaps   0;

Db    159 hvchnvhbnnhrnwayvrhdarrddvh 185
          ::|:  ::||  |   ::::|::::::
Cp     29 tcctatacNNNNNNNNgaataggaact 3


RESULT   18
ID   Q49264 standard; DNA; 4093 BP.
AC   Q49264;
DT   28-APR-1994  (first entry)
DE   ced-4.
KW   Long-distance homology; evolution; nematode;
KW   hybridisation; lower organism; structural homologue;
KW   Alzheimer's disease; cell death gene; PCR; polymerase chain reaction;
KW   ciona intestinalis; echinoderm; lamprey; puffer fish;
KW   mammal; probe; ds.
OS   Caenorhabditis briggsae.
FH   Key             Location/Qualifiers
FT   CDS             459..3246
FT   /*tag=  a
FT   /*product= ced-4_gene_product
FT   exon            459..908
FT   /*tag=  b
FT   exon            986..1081
FT   /*tag=  c
FT   exon            1383..1472
FT   /*tag=  d
FT   exon            1651..1716
FT   /*tag=  e
```

```
FT   exon            1834..2172
FT   /*tag=  f
FT   exon            2477..2752
FT   /*tag=  g
FT   exon            2802..2906
FT   /*tag=  h
FT   exon            3031..3246
FT   /*tag=  i
PN   WO9320237-A.
PD   14-OCT-1993.
PF   01-APR-1993; U03102.
PR   01-APR-1992; US-861458.
PA   (CAMB-) CAMBRIDGE NEUROSCIENCE INC.
PI   Johnson CD,  Marchionni MA;
DR   WPI; 93-336943/42.
DR   P-PSDB; R42742.
PT   Long-distance homology cloning of genes from lower organisms -
PT   used to identify DNA that codes for evolutionary conserved
PT   aminoacid sequences
PS   Disclosure; Fig 8; 188pp; English.
CC   The primers/probes (Q49266-Q49295) are used to isolate the ced-4
CC   gene from the nematode C. briggsae.
SQ   Sequence   4093 BP;    1226 A;    792 C;    726 G;    1346 T;

  Query Match           61.5%;  Score 16;  DB 9;  Length 4093;
  Best Local Similarity  61.8%;  Pred. No. 2.91e+00;
  Matches    21;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db   3228 gaagttcacatcccaatagctgtataagaatttc 3261
          |||||||  || |        ||||| ||| |||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   19
ID   Q67134 standard; DNA; 42 BP.
AC   Q67134;
DT   22-MAR-1995  (first entry)
DE   DNA primer used for construction of FRT containing vectors.
KW   DNA primer; FRT sequence; vector; maize; Zea mays; cereal; grass;
KW   protoplast; ss.
OS   Synthetic.
PN   WO9417176-A.
PD   04-AUG-1994.
PF   27-JAN-1994; U00927.
PR   29-JAN-1993; US-010997.
PA   (PURD ) PURDUE RES FOUND.
PI   Hodges TK,  Lyznik LA;
DR   WPI; 94-264090/32.
PT   DNA constructs - for creating transgenic eukaryotic cells
PS   Disclosure; Page 49; 79pp; English.
CC   This primer is used in the construction of FRT containing vectors
CC   which are used in the construction of trnagenic eukaryotic cells.
CC   This primer is annealed to another primer (Q67135) and incubated
CC   with T4 DNA-polymerase and each dNTP to form a complete FRT
CC   recombination site of 48 bp.
SQ   Sequence   42 BP;    12 A;    9 C;    7 G;    14 T;

  Query Match           57.7%;  Score 15;  DB 12;  Length 42;
  Best Local Similarity  64.0%;  Pred. No. 9.70e+00;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db      8 gaagttcctattccgaagttcctat 32
          ||||||||||||         |||
Qy      1 gaagttcctattcNNNNNNNNgtat 25
```



```
RESULT   20
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match           57.7%;  Score 15;  DB 2;  Length 1047;
  Best Local Similarity  21.2%;  Pred. No. 9.70e+00;
  Matches     7;  Conservative   13;  Mismatches  13;  Indels   0;  Gaps   0;
```

```
Db     83 savdnknyhdndnnngngcvynaasvarnashw 115
          :|:: : ::: : || |     || : : |:::
Cp     34 gaagttcctatacNNNNNNNNgaataggaactt 2
```

RESULT 21
ID Q56791 standard; cDNA; 1971 BP.
AC Q56791;
DT 07-OCT-1994 (first entry)
DE cDNA encoding receptor for C-terminus of beta-amyloid precursor.
KW Receptor; precursor protein; alzheimers disease; antibodies;
KW transgenic animal; diagnosis; detection; therapy; agonist;
KW antagonist; antisense; ribozyme; beta amyloid precursor protein;
KW C100-R; ss.
OS Rattus rattus.
FH Key Location/Qualifiers
FT CDS 1..1395
FT /*tag= a
FT /product= Beta amyloid precursor protein receptor.
PN WO9405811-A.
PD 17-MAR-1994.
PF 31-AUG-1993; U08229.
PR 31-AUG-1992; US-938184.
PR 30-AUG-1993; US-938184.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PA (MCLE-) MCLEAN HOSPITAL CORP.
PI Kozlowski MR, Manly SP, Neve RL;
DR WPI; 94-101215/12.
DR P-PSDB; R50951.
PT Cloning and expression of beta APP-C100 receptor - facilitating
PT study, diagnosis and therapy of Alzheimer's disease
PS Claim 8; Page 49-51; 78pp; English.
CC The cDNA encodes a receptor for the C-terminus of the beta amyloid
CC precursor protein (the C100-R) and so facilitates the elucidation of
CC the function of C100-R and its role in the development of Alzheimers
CC disease. It may be used in hybridisation assays of biopsies or
CC autopsies to diagnose abnormalities of C100-R expression. Antisense
CC or ribozyme molecules designed on the basis of the C100-R DNA
CC sequence may be utilised to block transcription and expression of the
CC C100-R gene. Antibodies specific for the C100-R may be used to
CC determine the pattern of receptor expression in biopsy tissue, or for
CC diagnostic imaging in vivo. Transgenic animals containing the C100-R
CC DNA as the transgene may be engineered to determine the in vivo
CC effects of the beta amyloid precursor protein-C100 agonists or
CC antagonists, or to profile other agents which are potentially
CC therapeutic for alzheimers disease.
SQ Sequence 1971 BP; 620 A; 440 C; 436 G; 475 T;

Query Match 57.7%; Score 15; DB 10; Length 1971;
Best Local Similarity 61.3%; Pred. No. 9.70e+00;
Matches 19; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Db   1499 gttctttaactttcaagtgaatagaaacttc 1529
          |||| |  ||       |||||| ||||||
Cp     31 gttcctatacNNNNNNNNgaataggaacttc 1
```

RESULT 22
ID Q71367 standard; DNA; 3249 BP.
AC Q71367;
DT 21-APR-1995 (first entry)
DE E.coli/S.cerevisiae shuttle vector pMTL8100.
KW Cassette; gene expression; promoter; recombinant protein;

KW fermentation; heterologous gene; clone; cloning; yeast; bacteria;
KW 2 mu plasmid; ds.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_signal 3003..3225
FT /*tag= a
FT /label= 2mu replication region.
FT misc_feature 2375..2666
FT /*tag= b
FT /label= STB locus.
FT CDS 461..1117
FT /*tag= c
FT /product= Chloramphenicol acetyltransferase
PN WO9419472-A.
PD 01-SEP-1994.
PF 25-FEB-1994; G00373.
PR 26-FEB-1993; GB-003988.
PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
PI Faulkner JDB, Minton NP;
DR WPI; 94-294335/36.
PT New promoter DNA with unique SspI site at gene start position -
PT esp modified yeast promoter, provides high level of recombinant
PT protein expression in bacteria and yeast
PS Example 3; Page 32-34; 48pp; English.
CC This shuttle vector has the replicative functions of an E.coli
CC plasmid as well as those of a S.cerevisiae plasmid. The vector was
CC constructed by isolating a 1.4kb RsaI fragment which encompassed the
CC origin of replication and STB locus of the 2mu plasmid, from plasmid
CC pVT100-U and inserting it into the unique EcoRV site of pMTLCJ.
CC Plasmid pMTLCJ was constructed essentially by cloning a 0.8 kb
CC BamHI fragment encoding chloramphenicol acetyltransferase (cat) from
CC plasmid pCM4 (Close and Rodriguez, 1982) into the BamHI site of
CC M13mp8. Single stranded DNA prepared from the resulting recombinant
CC was then used as a template in successive site directed mutagenesis
CC to eliminate restriction sites from the cat structural gene. Double
CC stranded DNA of the mutated M13 recombinant was then prepared and
CC the modified cat gene excised as a 0.8 kb BamHI fragment, which was
CC then blunt ended and ligated to a 1.1kb SspI/DraI fragment
CC encompassing the replication region of plasmid pMTL4 (Chambers et
CC al., 1988), to give pMTLCJ.
SQ Sequence 3249 BP; 882 A; 693 C; 743 G; 931 T;

Query Match 57.7%; Score 15; DB 12; Length 3249;
Best Local Similarity 64.0%; Pred. No. 9.70e+00;
Matches 16; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Db   116 ataggaacttcggaataggaacttc 140
            |||         ||||||||||||
Cp    25 atacNNNNNNNNgaataggaacttc 1
```

RESULT 23
ID Q71366 standard; DNA; 3400 BP.
AC Q71366;
DT 21-APR-1995 (first entry)
DE E.coli/S.cerevisiae shuttle vector pMTL8000.
KW Cassette; gene expression; promoter; recombinant protein;
KW fermentation; heterologous gene; clone; cloning; yeast; bacteria;
KW 2 mu plasmid; ds.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_signal 3154..3376
FT /*tag= a
FT /label= 2mu replication region.

```
FT   misc_feature    2526..2817
FT   /*tag=  b
FT   /label= STB locus.
FT   CDS             444..1304
FT   /*tag=  c
FT   /product= Beta lactamase.
PN   WO9419472-A.
PD   01-SEP-1994.
PF   25-FEB-1994; G00373.
PR   26-FEB-1993; GB-003988.
PA   (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
PI   Faulkner JDB,  Minton NP;
DR   WPI; 94-294335/36.
PT   New promoter DNA with unique SspI site at gene start position -
PT   esp modified yeast promoter, provides high level of recombinant
PT   protein expression in bacteria and yeast
PS   Example 3; Page 30-32; 48pp; English.
CC   This shuttle vector has the replicative functions of an E.coli
CC   plasmid as well as those of a S.cerevisiae plasmid.  The vector was
CC   constructed by isolating a 1.4kb RsaI fragment which encompassed the
CC   origin of replication and STB locus of the 2mu plasmid, from plasmid
CC   pVT100-U and inserting it into the unique EcoRV site of pMTLJ.
CC   Plasmid pMTLJ was derived from pMTL4 (Chambers at al., 1988), by
CC   eliminating the SspI restriction site using the plasmid site
CC   directed mutagenesis method.
SQ   Sequence   3400 BP;    917 A;    738 C;    787 G;    958 T;

  Query Match           57.7%;  Score 15;  DB 12;  Length 3400;
  Best Local Similarity  64.0%;  Pred. No. 9.70e+00;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    116 ataggaacttcggaataggaacttc 140
          |||         ||||||||||||
Cp     25 atacNNNNNNNNgaataggaacttc 1


RESULT   24
ID   Q77789 standard; DNA; 5211 BP.
AC   Q77789;
DT   26-JUN-1995  (first entry)
DE   Pre-pro-cobra C3 coding sequence.
KW   Cobra; C3; third component of complement; human; mouse; rat;
KW   X. laevis; pre-pro molecule; beta chain; alpha chain; codon usage;
KW   G+C content; immune response; host defence; ss.
OS   Naja naja.
FH   Key             Location/Qualifiers
FT   CDS             9..4964
FT   /*tag=  a
FT   /product= Pre-pro-cobra C3
FT   sig_peptide     9..74
FT   /*tag=  b
FT   mat_peptide     75..4961
FT   /*tag=  c
FT   misc_difference 480..482
FT   /*tag=  d
FT   /codon= seq:GAA, aa:Asp
FT   misc_difference 483..485
FT   /*tag=  e
FT   /codon= seq:CAA, aa:Lys
PN   WO9423024-A.
PD   13-OCT-1994.
PF   07-APR-1994; U03441.
PR   07-APR-1993; US-043747.
PA   (GEOU ) UNIV GEORGETOWN.
```



```
PI   Bredehorst R,  Fritzinger DC,  Vogel C;
DR   WPI; 94-333186/41.
DR   P-PSDB; R63222.
PT   DNA encoding cobra C3, CVF 1 and CVF 2 - which are used in the
PT   treatment of cancer
PS   Claim 1; Fig 2A-2L; 155pp; English.
CC   This sequence encodes the cobra C3 (third component of complement).
CC   The cDNA sequence of cobra C3 shows a high sequence homology with C3
CC   molecules from human, mouse, rat and X. laevis.  Cobra C3 is
CC   synthesised as a pre-pro molecule that is subsequently processed
CC   into the mature two-chain protein by removing the signal peptide and
CC   the four Arg residues between the beta and alpha chain.  The alpha
CC   chain comprises 992 amino acids and the beta chain comprises 633
CC   residues, being 12 residues shorter than the human beta chain.  Cobra
CC   C3 has a different codon usage compared to mammalian C3 mRNAs.  The
CC   G+C content of all known mammalian C3 mRNAs is more than 53%.  The
CC   G+C content of cobra C3 mRNA is significantly lower at 43%.  The
CC   significance of this difference is not known.  C3 is thought to have
CC   important functions in the immune response and host defence.
SQ   Sequence   5211 BP;    1612 A;    1042 C;    1201 G;    1356 T;

  Query Match           57.7%;  Score 15;  DB 13;  Length 5211;
  Best Local Similarity  60.6%;  Pred. No. 9.70e+00;
  Matches    20;  Conservative    0;  Mismatches   13;  Indels   0;  Gaps   0;

Db    711 aagtttctttacattgatgggaataaaaatttc 743
          ||||| || |||        |||||  || |||
Cp     33 aagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   25
ID   Q39050 standard; DNA; 6824 BP.
AC   Q39050;
DT   28-JUL-1993  (first entry)
DE   K.lactis/S. cerevisae genetic vector.
KW   Genetic; vector; integration; Kluyeromyces lactis; 255 ribosomal DNA;
KW   Saccharomyces cerevisiae; E. coli; domain; yeast; plasmid; promoter;
KW   expression cassette; HIS3; marker; transformant; human; lysozyme; HLZ;
KW   GAL7; signal sequence; killer toxin; transcription termination signal;
KW   FLP; 2 micron plasmid; ss.
OS   Synthetic.
PN   EP-537456-A.
PD   21-APR-1993.
PF   31-AUG-1992; 114838.
PR   04-SEP-1991; IT-MI2349.
PA   (ISTS ) SCLAVO SPA.
PI   Galeotti CL,  Gallo E,  Riccio ML,  Rossolini GM,  Thaller MC;
DR   WPI; 93-127394/16.
PT   Vector for Kluyeromyces lactis and Saccharomyces cerevisiae -
PT   which allows stable multiple integration of DNA for prodn. of
PT   heterologous proteins
PS   Claim 1; Fig 1; 26pp; English.
CC   This sequence represents a genetic vector which allows the stable
CC   multiple integration of DNA sequences into the genome of Kluyeromyces
CC   lactis and Saccharomyces cerevisiae.  This sequence can be used in an
CC   integrating vector which comprises a region necessary for the stable
CC   maintenance of the plasmid in E. coli and a domain which acts as an
CC   integrating unit consisting of two not contiguous sequences of the 255
CC   ribosomal DNA from S. cerevisiae, flanking a genetic marker suitable
CC   for selection of the yeast transformants in which the integration
CC   event has occured.  Other DNA sequences may be introduced into the
CC   integration plasmid, such as expression cassettes.  The gene HIS3
CC   from K. lactis and S. cerevisiae is pref. used as a genetic marker
CC   for the selection of transformants and an expression cassette for the
```

```
CC   production and secretion into the culture medium of human lysozyme.
CC   This complete transformation vector is 7850 bp long and includes the
CC   integration vector of the invention and an expression cassette
CC   comprising the K. lactis GAL7 promoter, the signal sequence of the K.
CC   lactis killer toxin, the cDNA encoding the ripe form of human lysozyme
CC   (HLZ) and the transcription termination signal FLP of the 2 micron
CC   plasmid from S. cerevisiae.
SQ   Sequence   6824 BP;    1815 A;    1521 C;    1726 G;    1762 T;

  Query Match           57.7%;  Score 15;  DB 7;  Length 6824;
  Best Local Similarity  64.0%;  Pred. No. 9.70e+00;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db   2506 gaagttcctattccgaagttcctat 2530
          ||||||||||||         |||
Qy      1 gaagttcctattcNNNNNNNNgtat 25


RESULT   26
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990  (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS   Escherichia coli.
FH   Key             Location/Qualifiers
FT   misc_feature    19..69
FT   /*tag=  a
FT   /function=multiple cloning site
FT   primer_bind     187..204
FT   /*tag=  b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI; 88-279927/40.
PT   Introducing random point mutations into nucleic acods -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; p; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence   204 BP;   21 A;   47 C;   17 G;   11 T;   108 Others;

  Query Match           53.8%;  Score 14;  DB 1;  Length 204;
  Best Local Similarity  18.5%;  Pred. No. 3.13e+01;
  Matches     5;  Conservative   14;  Mismatches   8;  Indels   0;  Gaps   0;

Db    159 hvchnvhbnnhrnwayvrhdarrddvh 185
          ::|:  ::||  |   : ::|::::::
Qy      6 tcctattcNNNNNNNNgtataggaact 32
```



```
RESULT   27
ID   N50034 standard; DNA; 498 BP.
AC   N50034;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 485.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..498
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50033.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 21, page 43; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   498 BP;    112 A;    30 C;    68 G;    77 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 498;
  Best Local Similarity  42.4%;  Pred. No. 3.13e+01;
  Matches    14;  Conservative    6;  Mismatches  13;  Indels   0;  Gaps   0;

Db    116 arathccnatggaratgacngaraargarttyc 148
          |: |:|| ||        |||: |:||: |:|
Cp     33 aagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   28
ID   N50025 standard; DNA; 501 BP.
AC   N50025;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 418.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50024.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
```

```
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2c, page 34; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    112 A;    30 C;    69 G;    85 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   29
ID   N50023 standard; DNA; 501 BP.
AC   N50023;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 416.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50022.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2a, page 32; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    107 A;    31 C;    69 G;    80 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   30
ID   N50026 standard; DNA; 501 BP.
AC   N50026;
DT   04-SEP-1991  (first entry)
```



```
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 430.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50025.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2d, page 35; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    108 A;    31 C;    70 G;    81 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   31
ID   N50031 standard; DNA; 501 BP.
AC   N50031;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 448.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50030.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2i, page 40; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
```

```
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    110 A;    30 C;    69 G;    80 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   32
ID   N50029 standard; DNA; 501 BP.
AC   N50029;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 446.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50028.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2g, page 38; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    112 A;    31 C;    69 G;    79 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   33
ID   N50027 standard; DNA; 501 BP.
AC   N50027;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 444.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
```



```
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50026.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2e, page 36; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    112 A;    31 C;    67 G;    80 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   34
ID   N50032 standard; DNA; 501 BP.
AC   N50032;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 449.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50031.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2j, page 41; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    108 A;    30 C;    72 G;    79 T;
```

```
  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   35
ID   N50024 standard; DNA; 501 BP.
AC   N50024;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 417.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50023.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2b, page 33; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    110 A;    32 C;    66 G;    81 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   36
ID   N50033 standard; DNA; 501 BP.
AC   N50033;
DT   04-SEP-1991  (first entry)
DE   Sequence encoding new modified human beta interferon polypeptides
DE   IFNX 456.
KW   Antiviral; cell growth regulator; immune system regulator;
KW   antiproliferative; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1..501
FT   /*tag=  a
PN   EP-163993-A.
PD   11-DEC-1985.
```



```
PF   17-MAY-1985; 105750.
PR   17-MAY-1984; GB-012564.
PA   (SEAR ) SEARLE G D & CO.
PI   Bell LD, Boseley PG, Porter AG;
DR   WPI; 85-311944/50.
DR   P-PSDB; P50032.
PT   New modified human beta interferon polypeptide(s) - prepd. by
PT   plasmid transformed bacteria, with improved antiviral,
PT   anti-proliferative and immune regulating actions
PS   Claim 28; Chart 2k, page 42; 71pp; English.
CC   Compared with interferon beta prepd. by recombinant methods, the
CC   INFs of the invention are more active and have different affinities
CC   for cell surface receptors (allowing selective targetting); they
CC   have higher therapeutic index; improved stability against microbial
CC   breakdown during synthesis; and better in vivo solubility and
CC   stability.  They are also easier to recover from incubation mixts.
SQ   Sequence   501 BP;    111 A;    31 C;    68 G;    80 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 501;
  Best Local Similarity  22.6%;  Pred. No. 3.13e+01;
  Matches     7;  Conservative    9;  Mismatches  15;  Indels   0;  Gaps   0;

Db     15 bytbggnttyytbcarmgdwnnwnnaaytty 45
          ::|:   ||:        ::  :  ||:||:
Qy      4 gttcctattcNNNNNNNNgtataggaacttc 34


RESULT   37
ID   Q25420 standard; DNA; 1561 BP.
AC   Q25420;
DT   30-NOV-1992  (first entry)
DE   Encodes human liver cysteine dioxygenase.
KW   cystine; genetic diagnosis; cystine urine diseases; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             230..830
PN   J04131083-A.
PD   01-MAY-1992.
PF   20-SEP-1990; 251647.
PR   20-SEP-1990; JP-251647.
PA   (AJIN ) AJINOMOTO KK.
DR   WPI; 92-197392/24.
DR   P-PSDB; R24407.
PT   Human liver cysteine di:oxygenase and cDNA used for its encoding
PT   - is used for diagnosis and treatment of cystine-associated
PT   urinary diseases
PS   Claim 3; Fig 2; 9pp; Japanese.
CC   This sequence encodes human cysteine dioxygenase. A cDNA library was
CC   prepared using polyA+ RNA separated from human non-cancer liver
CC   tissue from a liver cancer patient. The library was screened with
CC   rat cDNA clone RL-39(10) as probe. One clone (I)-1 was purified, and
CC   sequenced. The sequence is expected to be useful for genetic
CC   diagnosis and treatment, esp. for the treatment of cystine urine
CC   diseases, and cystine diseases.
SQ   Sequence   1561 BP;    474 A;    342 C;    337 G;    408 T;

  Query Match           53.8%;  Score 14;  DB 4;  Length 1561;
  Best Local Similarity  59.4%;  Pred. No. 3.13e+01;
  Matches    19;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db   1327 aaatttcttttctatgaataatagaggaactt 1358
          || || || |||         || ||||||||
Qy      2 aagttcctattcNNNNNNNNgtataggaactt 33
```

```
RESULT   38
ID   Q47839 standard; cDNA; 1997 BP.
AC   Q47839;
DT   16-MAR-1994  (first entry)
DE   Human interleukin 9 receptor clone ph9RA6.
KW   Interleukin 9 receptor; IL-9; antibodies; therapy; probe; agonist;
KW   antagonist; ss
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             188..1798
FT   /*tag=  a
FT   /product= Interleukin 9 receptor.
PN   WO9318047-A.
PD   16-SEP-1993.
PF   25-FEB-1993; U01720.
PR   09-MAR-1992; US-847347.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Druez C,  Renauld J,  Van Snick J;
DR   WPI; 93-303390/38.
PT   Nucleic acid encoding interleukin-9 receptor - used to produce
PT   reagents used in diagnosis and therapy involving interleukin 9R
PS   Claim 6; Page 18; 30pp; English.
CC   The interleukin (IL) 9 receptor nucleic acid sequence can be used to
CC   produce IL-9 receptor or as probes for cells which respond to the
CC   cytokine.  The complementary sequences can be used to inhibit the
CC   expression of the IL-9 receptor protein and to probe for the IL-9
CC   coding sequences.  Transfected cell lines can be used to screen for
CC   IL-9 receptor agonists and antagonists.  Antibodies directed against
CC   the IL-9 receptor can be used therapeutically to block IL-9 binding
CC   to the receptor and for qualitative and quantative measurement of
CC   IL-9 receptor levels.
SQ   Sequence   1997 BP;    388 A;    612 C;    593 G;    404 T;

  Query Match           53.8%;  Score 14;  DB 8;  Length 1997;
  Best Local Similarity  60.7%;  Pred. No. 3.13e+01;
  Matches    17;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db   1171 cctctacagtgtacacaatgggaacttc 1198
            ||| |||         ||| ||||||||
Cp     28 cctatacNNNNNNNNgaataggaacttc 1


RESULT   39
ID   Q71367 standard; DNA; 3249 BP.
AC   Q71367;
DT   21-APR-1995  (first entry)
DE   E.coli/S.cerevisiae shuttle vector pMTL8100.
KW   Cassette; gene expression; promoter; recombinant protein;
KW   fermentation; heterologous gene; clone; cloning; yeast; bacteria;
KW   2 mu plasmid; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_signal     3003..3225
FT   /*tag=  a
FT   /label= 2mu replication region.
FT   misc_feature    2375..2666
FT   /*tag=  b
FT   /label= STB locus.
FT   CDS             461..1117
FT   /*tag=  c
FT   /product= Chloramphenicol acetyltransferase
PN   WO9419472-A.
PD   01-SEP-1994.
PF   25-FEB-1994; G00373.
```



```
PR   26-FEB-1993; GB-003988.
PA   (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
PI   Faulkner JDB,  Minton NP;
DR   WPI; 94-294335/36.
PT   New promoter DNA with unique SspI site at gene start position -
PT   esp modified yeast promoter, provides high level of recombinant
PT   protein expression in bacteria and yeast
PS   Example 3; Page 32-34; 48pp; English.
CC   This shuttle vector has the replicative functions of an E.coli
CC   plasmid as well as those of a S.cerevisiae plasmid.  The vector was
CC   constructed by isolating a 1.4kb RsaI fragment which encompassed the
CC   origin of replication and STB locus of the 2mu plasmid, from plasmid
CC   pVT100-U and inserting it into the unique EcoRV site of pMTLCJ.
CC   Plasmid pMTLCJ was constructed essentially by  cloning a 0.8 kb
CC   BamHI fragment encoding chloramphenicol acetyltransferase (cat) from
CC   plasmid pCM4 (Close and Rodriguez, 1982) into the BamHI site of
CC   M13mp8.  Single stranded DNA prepared from the resulting recombinant
CC   was then used as a template in successive site directed mutagenesis
CC   to eliminate restriction sites from the cat structural gene. Double
CC   stranded DNA of the mutated M13 recombinant was then prepared and
CC   the modified cat gene excised as a 0.8 kb BamHI fragment, which was
CC   then blunt ended and ligated to a 1.1kb SspI/DraI fragment
CC   encompassing the replication region of plasmid pMTL4 (Chambers et
CC   al., 1988), to give pMTLCJ.
SQ   Sequence   3249 BP;    882 A;    693 C;    743 G;    931 T;

  Query Match           53.8%;  Score 14;  DB 12;  Length 3249;
  Best Local Similarity  62.5%;  Pred. No. 3.13e+01;
  Matches    15;  Conservative    0;  Mismatches  9;  Indels   0;  Gaps   0;

Db   103 ttctagctagagaataggaacttc 126
           |||        | ||||||||||
Qy    11 ttcNNNNNNNNgtataggaacttc 34


RESULT   40
ID   Q71366 standard; DNA; 3400 BP.
AC   Q71366;
DT   21-APR-1995  (first entry)
DE   E.coli/S.cerevisiae shuttle vector pMTL8000.
KW   Cassette; gene expression; promoter; recombinant protein;
KW   fermentation; heterologous gene; clone; cloning; yeast; bacteria;
KW   2 mu plasmid; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_signal     3154..3376
FT   /*tag=  a
FT   /label= 2mu replication region.
FT   misc_feature    2526..2817
FT   /*tag=  b
FT   /label= STB locus.
FT   CDS             444..1304
FT   /*tag=  c
FT   /product= Beta lactamase.
PN   WO9419472-A.
PD   01-SEP-1994.
PF   25-FEB-1994; G00373.
PR   26-FEB-1993; GB-003988.
PA   (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
PI   Faulkner JDB,  Minton NP;
DR   WPI; 94-294335/36.
PT   New promoter DNA with unique SspI site at gene start position -
PT   esp modified yeast promoter, provides high level of recombinant
PT   protein expression in bacteria and yeast
```

```
PS   Example 3; Page 30-32; 48pp; English.
CC   This shuttle vector has the replicative functions of an E.coli
CC   plasmid as well as those of a S.cerevisiae plasmid.  The vector was
CC   constructed by isolating a 1.4kb RsaI fragment which encompassed the
CC   origin of replication and STB locus of the 2mu plasmid, from plasmid
CC   pVT100-U and inserting it into the unique EcoRV site of pMTLJ.
CC   Plasmid pMTLJ was derived from pMTL4 (Chambers at al., 1988), by
CC   eliminating the SspI restriction site using the plasmid site
CC   directed mutagenesis method.
SQ   Sequence   3400 BP;    917 A;    738 C;    787 G;    958 T;

  Query Match           53.8%;  Score 14;  DB 12;  Length 3400;
  Best Local Similarity  62.5%;  Pred. No. 3.13e+01;
  Matches    15;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    103 ttctagctagagaataggaacttc 126
          |||        | ||||||||||
Qy     11 ttcNNNNNNNNgtataggaacttc 34


RESULT   41
ID   Q49264 standard; DNA; 4093 BP.
AC   Q49264;
DT   28-APR-1994  (first entry)
DE   ced-4.
KW   Long-distance homology; evolution; nematode;
KW   hybridisation; lower organism; structural homologue;
KW   Alzheimer's disease; cell death gene; PCR; polymerase chain reaction;
KW   ciona intestinalis; echinoderm; lamprey; puffer fish;
KW   mammal; probe; ds.
OS   Caenorhabditis briggsae.
FH   Key             Location/Qualifiers
FT   CDS             459..3246
FT   /*tag= a
FT   /*product= ced-4_gene_product
FT   exon            459..908
FT   /*tag= b
FT   exon            986..1081
FT   /*tag= c
FT   exon            1383..1472
FT   /*tag= d
FT   exon            1651..1716
FT   /*tag= e
FT   exon            1834..2172
FT   /*tag= f
FT   exon            2477..2752
FT   /*tag= g
FT   exon            2802..2906
FT   /*tag= h
FT   exon            3031..3246
FT   /*tag= i
PN   WO9320237-A.
PD   14-OCT-1993.
PF   01-APR-1993; U03102.
PR   01-APR-1992; US-861458.
PA   (CAMB-) CAMBRIDGE NEUROSCIENCE INC.
PI   Johnson CD,  Marchionni MA;
DR   WPI; 93-336943/42.
DR   P-PSDB; R42742.
PT   Long-distance homology cloning of genes from lower organisms -
PT   used to identify DNA that codes for evolutionary conserved
PT   aminoacid sequences
PS   Disclosure; Fig 8; 188pp; English.
CC   The primers/probes (Q49266-Q49295) are used to isolate the ced-4
```



```
CC   gene from the nematode C. briggsae.
SQ   Sequence   4093 BP;    1226 A;    792 C;    726 G;    1346 T;

  Query Match           53.8%;  Score 14;  DB 9;  Length 4093;
  Best Local Similarity  58.8%;  Pred. No. 3.13e+01;
  Matches    20;  Conservative    0;  Mismatches  14;  Indels   0;  Gaps   0;

Db   3228 gaagttcacatcccaatagctgtataagaatttc 3261
          |||||||  || |        | ||| ||| |||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   42
ID   Q24802 standard; DNA; 10097 BP.
AC   Q24802;
DT   06-JUL-1992  (first entry)
DE   SIVmac239 nef-deletion.
KW   Macaque; monkey; polymerase chain reaction;
KW   PCR; site-directed mutagenesis; retrovirus; null mutation; ss.
OS   Simian immunodeficiency virus.
FH   Key             Location/Qualifiers
FT   repeat_region   1..818
FT   /*tag= a
FT   /rpt_type= TERMINAL
FT   /note= "i.e. 5' LTR"
FT   repeat_unit     1..517
FT   /*tag= b
FT   /rpt_type= OTHER
FT   /note= "U3"
FT   repeat_unit     518..600
FT   /*tag= c
FT   /rpt_type= OTHER
FT   /note= "R"
FT   repeat_unit     601..818
FT   /*tag= d
FT   /rpt_type= OTHER
FT   /note= "U5"
FT   primer_bind     822..849
FT   /*tag= e
FT   /standard_name= tRNA_PBS
FT   CDS             1053..2585
FT   /*tag= f
FT   /product= gag
FT   CDS             2228..5410
FT   /*tag= g
FT   /product= pol
FT   CDS             5340..5984
FT   /*tag= h
FT   /product= vif
FT   CDS             5812..6150
FT   /*tag= i
FT   /product= vpx
FT   CDS             6051..6456
FT   /*tag= j
FT   /product= vpr
FT   exon            6302..6597
FT   /*tag= k
FT   /product= tat
FT   /note= "full-length product obtained by splicing"
FT   exon            6528..6597
FT   /*tag= l
FT   /product= rev
FT   /note= "full-length product obtained by splicing"
FT   CDS             6604..9243
```

```
FT   /*tag= m
FT   /product= env
FT   exon            8803..8902
FT   /*tag= n
FT   /product= tat
FT   /note= "see above"
FT   exon            8803..9059
FT   /*tag= o
FT   /product= rev
FT   /note= "see above"
FT   CDS             9077..9686
FT   /*tag= p
FT   /product= nef
FT   repeat_region   9280..10097
FT   /*tag= q
FT   /rpt_type= TERMINAL
FT   /note= "i.e. 3' LTR"
FT   misc_signal     415..424
FT   /*tag= r
FT   /standard_name= NF_Kappa_B
FT   GC_signal       429..438
FT   /*tag= s
FT   /standard_name= Sp1_binding_site
FT   GC_signal       440..449
FT   /*tag= t
FT   /standard_name= Sp1_binding_site
FT   GC_signal       451..460
FT   /*tag= u
FT   /standard_name= Sp1_binding_site
FT   GC_signal       462..471
FT   /*tag= v
FT   /standard_name= Sp1_binding_site
FT   TATA_signal     488..494
FT   /*tag= w
FT   polyA_signal    9950..9955
FT   /*tag= x
PN   WO9200987-A.
PD   23-JAN-1992.
PF   10-JUL-1991; U04884.
PR   12-JUL-1990; US-551945.
PA   (HARD ) HARVARD COLLEGE.
PI   Desrosiers RC.
DR   WPI; 92-056816/07.
DR   P-PSDB; R22365-R22371, R24126-7.
PT   Primate lentivirus vaccine protecting against AIDS - and primate
PT   lentiviruses and their DNA clones contg. null mutations, useful for
PT   producing vaccine
PS   Claim 1; Fig 1; 51pp; English.
CC   Parental virus SIVmac239 was isolated from a macaque monkey (see
CC   Q22487). An SstI fragment of p239SpE3' contg. the C-terminus of
CC   the nef gene was cloned into M13 and subjected to site-directed
CC   mutagenesis. A 73mer primer was used which was complementary to
CC   bases 9215 through 9250 and 9433 through 9469. Since bases 9251
CC   through 9432 were not included in the primer, their complement was
CC   not included in the newly synthesised negative strand. Successful
CC   deletion was confirmed by DNA sequencing. The fragment contg. the
CC   deletion was cloned into p239SpE3' to yield p239S-E3' (nef-deletion)
CC   which was digested with SphI and ligated to SphI-cut p239SpSp5'.
CC   The ligation product was used to directly transfect cultured cells.
CC   See also Q21075-8.
SQ   Sequence   10097 BP;    3387 A;    1911 C;    2527 G;    2272 T;

  Query Match           53.8%;  Score 14;  DB 3;  Length 10097;
  Best Local Similarity 61.5%;  Pred. No. 3.13e+01;
```



```
  Matches   16;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   7796 agttcctctactgtaaaatgaattgg 7821
          ||||||| |||        |||| ||
Cp     32 agttcctatacNNNNNNNNgaatagg 7


RESULT   43
ID   Q22487 standard; DNA; 10279 BP.
AC   Q22487;
DT   06-JUL-1992  (first entry)
DE   SIVmac239 proviral genome.
KW   Macaque; monkey; mac239; polymerase chain reaction;
KW   PCR; site-directed mutagenesis; retrovirus; ss.
OS   Simian immunodeficiency virus.
FH   Key             Location/Qualifiers
FT   repeat_region   1..818
FT   /*tag= a
FT   /rpt_type= TERMINAL
FT   /note= "i.e. 5' LTR"
FT   repeat_unit     1..517
FT   /*tag= b
FT   /rpt_type= OTHER
FT   /note= "U3"
FT   repeat_unit     518..600
FT   /*tag= c
FT   /rpt_type= OTHER
FT   /note= "R"
FT   repeat_unit     601..818
FT   /*tag= d
FT   /rpt_type= OTHER
FT   /note= "U5"
FT   primer_bind     822..849
FT   /*tag= e
FT   /standard_name= tRNA_PBS
FT   CDS             1053..2585
FT   /*tag= f
FT   /product= gag
FT   CDS             2228..5410
FT   /*tag= g
FT   /product= pol
FT   CDS             5340..5984
FT   /*tag= h
FT   /product= vif
FT   CDS             5812..6150
FT   /*tag= i
FT   /product= vpx
FT   CDS             6051..6456
FT   /*tag= j
FT   /product= vpr
FT   exon            6302..6597
FT   /*tag= k
FT   /product= tat
FT   /note= "full-length product obtained by splicing"
FT   exon            6528..6597
FT   /*tag= l
FT   /product= rev
FT   /note= "full-length product obtained by splicing"
FT   CDS             6604..9243
FT   /*tag= m
FT   /product= env
FT   exon            8803..8902
FT   /*tag= n
FT   /product= tat
```

```
FT   /note= "see above"
FT   exon            8803..9059
FT   /*tag= o
FT   /product= rev
FT   /note= "see above"
FT   CDS             9077..9868
FT   /*tag= p
FT   /product= nef
FT   repeat_region   9462..10279
FT   /*tag= q
FT   /rpt_type= TERMINAL
FT   /note= "i.e. 3' LTR"
FT   misc_signal     415..424
FT   /*tag= r
FT   /standard_name= NF_Kappa_B
FT   GC_signal       429..438
FT   /*tag= s
FT   /standard_name= Sp1_binding_site
FT   GC_signal       440..449
FT   /*tag= t
FT   /standard_name= Sp1_binding_site
FT   GC_signal       451..460
FT   /*tag= u
FT   /standard_name= Sp1_binding_site
FT   GC_signal       462..471
FT   /*tag= v
FT   /standard_name= Sp1_binding_site
FT   TATA_signal     488..494
FT   /*tag= w
FT   polyA_signal    10132..10137
FT   /*tag= x
PN   WO9200987-A.
PD   23-JAN-1992.
PF   10-JUL-1991; U04884.
PR   12-JUL-1990; US-551945.
PA   (HARD ) HARVARD COLLEGE.
PI   Desrosiers RC.
DR   WPI; 92-056816/07.
DR   P-PSDB; R22365-R22372, R24126.
PT   Primate lentivirus vaccine protecting against AIDS - and primate
PT   lentiviruses and their DNA clones contg. null mutations, useful for
PT   producing vaccine
PS   Disclosure; Fig 1; 51pp; English.
CC   Cell-free serum samples from a macaque monkey exhibiting symptoms
CC   characteristic of SIV infection were co-cultivated with HuT-78
CC   cells. Infectious SIVmac239 virus was identified in the cell
CC   supernatant. Total cell DNA was prepared from SIVmac239-infected
CC   cells and digested with EcoRI. An EMBL-4 library was constructed
CC   from 10-20kb EcoRI fragments (EcoRI is a non-cutter of SIVmac239).
CC   The library was screened with pK2 BamA as probe and a full-length
CC   molecular clone was isolated and sequenced. Then, EMBL-SIVmac239
CC   was digested with SphI and a 6706bp fragment, contg. the SphI site
CC   in the left flanking cellular DNA sequence to viral nucleotide no.
CC   6451, was inserted in vector pBS(+) to produce subclone p239SpSp5'.
CC   In a separate reaction, EMBL-SIVmac239 was digested with EcoRI and
CC   SphI. A 6361bp fragment from viral nucleotide 6452 to the EcoRI
CC   site in the right flanking cellular sequence, was inserted in
CC   pBS(-) to produce subclone p239SpE3'. These subclones were used to
CC   generate the full-length genomic sequence and to produce the
CC   preferred null-mutations of the invention. See Q24802 for
CC   nef-deletion mutant. See also Q21075-8.
SQ   Sequence   10279 BP;    3465 A;    1936 C;    2569 G;    2309 T;

  Query Match          53.8%;  Score 14;  DB 3;  Length 10279;
```



```
  Best Local Similarity  61.5%;  Pred. No. 3.13e+01;
  Matches    16;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db   7796 agttcctctactgtaaaatgaattgg 7821
          ||||||| |||        |||| ||
Cp     32 agttcctatacNNNNNNNNgaatagg 7


RESULT   44
ID   Q54676 standard; DNA; 12151 BP.
AC   Q54676;
DT   03-AUG-1994  (first entry)
DE   Rice starch branching enzyme gene.
KW   Rice; starch; transit peptide; pectin; cereal; aminopectin; seeds;
KW   reverse transcriptase; plaques; ds.
OS   Oryza sativa.
FH   Key             Location/Qualifiers
FT   promoter        636..3351
FT   /*tag= a
FT   misc_binding    3164..3172
FT   /*tag= b
FT   CAAT_signal     3221..3225
FT   /*tag= c
FT   TATA_signal     3291..3296
FT   /*tag= d
FT   transit_peptide 3360..3443
FT   /*tag= e
FT   transit_peptide 3546..3608
FT   /*tag= f
FT   transit_peptide 5821..5853
FT   /*tag= g
FT   mat_peptide     5854..6028
FT   /*tag= h
FT   mat_peptide     6144..6231
FT   /*tag= i
FT   mat_peptide     6648..6917
FT   /*tag= j
FT   mat_peptide     7026..7932
FT   /*tag= k
FT   mat_peptide     8245..8361
FT   /*tag= l
FT   mat_peptide     8519..8581
FT   /*tag= m
FT   mat_peptide     9019..9126
FT   /*tag= n
FT   mat_peptide     9595..9696
FT   /*tag= o
FT   mat_peptide     9862..9929
FT   /*tag= p
FT   mat_peptide     10011..10091
FT   /*tag= q
FT   mat_peptide     10210..10326
FT   /*tag= r
FT   mat_peptide     10409..10609
FT   /*tag= s
FT   exon            3352..3443
FT   /*tag= t
FT   intron          3444..3545
FT   /*tag= u
FT   exon            3546..3608
FT   /*tag= v
FT   intron          3609..5820
FT   /*tag= w
FT   exon            5821..6028
```

```
FT   /*tag= x
FT   intron          6029..6143
FT   /*tag= y
FT   exon            6144..6213
FT   /*tag= z
FT   intron          6214..6647
FT   /*tag= aa
FT   exon            6648..6917
FT   /*tag= ab
FT   intron          6918..7025
FT   /*tag= ac
FT   exon            7026..7932
FT   /*tag= ad
FT   intron          7933..8244
FT   /*tag= ae
FT   exon            8245..8361
FT   /*tag= af
FT   intron          8362..8518
FT   /*tag= ag
FT   exon            8519..8581
FT   /*tag= ah
FT   intron          8582..9018
FT   /*tag= ai
FT   exon            9019..9126
FT   /*tag= aj
FT   intron          9127..9694
FT   /*tag= ak
FT   exon            9595..9696
FT   /*tag= al
FT   intron          9697..9861
FT   /*tag= am
FT   exon            9862..9929
FT   /*tag= an
FT   intron          9830..10010
FT   /*tag= ao
FT   exon            10011..10091
FT   /*tag= ap
FT   intron          10092..10209
FT   /*tag= aq
FT   exon            10210..10326
FT   /*tag= ar
FT   intron          10327..10408
FT   /*tag= as
FT   exon            10409..10865
FT   /*tag= at
FT   3'UTR           10610..10865
FT   /*tag= au
FT   polyA_signal    10814..10817
FT   /*tag= av
FT   polyA_signal    10833..10842
FT   /*tag= aw
FT   polyA_site      10865..10866
FT   /*tag= ax
PN   J05317057-A.
PD   03-DEC-1993.
PF   30-MAR-1992; 102499.
PR   20-SEP-1991; JP-268617.
PA   (MITS-) MITSUI GYOSAI SHOKUBUTU BIO KENKYUSHO KK.
DR   WPI; 94-011022/02.
DR   P-PSDB; R47469.
PT   Gene CDNA for rice starch branching enzyme for varied amino
PT   pectin in cereal - comprises structural gene specified by basic
PT   sequence introduced in rice plant for improved taste, for DNA
PT   fragment originated from rice genome contg. gene
```



```
PS   Claim 4; Page 10-11; 21pp; Japanese.
CC   The sequence shows a gene encoding a branching enzyme of rice starch.
CC   The enzyme can be used to modify aminopectin content of starch in
CC   cereal particles by introducing the basic sequence into a rice plant.
CC   This process can be used to improve the taste of the rice.
SQ   Sequence   12151 BP;    3273 A;    2479 C;    2506 G;    3890 T;

  Query Match           53.8%;  Score 14;  DB 10;  Length 12151;
  Best Local Similarity  60.0%;  Pred. No. 3.13e+01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   9848 aagttccttttcaggttattgtctttgaac 9877
          |||||||| |||        || |  ||||
Qy      2 aagttcctattcNNNNNNNNgtataggaac 31


RESULT   45
ID   Q62137 standard; cDNA to mRNA; 12151 BP.
AC   Q62137;
DT   07-MAR-1995  (first entry)
DE   Rice starch branching enzyme gene.
KW   Rice starch branching enzyme; oryza sativa; amylopectin; albumen;
KW   starch; ss.
OS   Oryza sativa.
FH   Key             Location/Qualifiers
FT   promoter        636..3351
FT   /*tag= a
FT   misc_binding    3164..3172
FT   /*tag= b
FT   CAAT_signal     3221..3225
FT   /*tag= c
FT   TATA_signal     3291..3296
FT   /*tag= d
FT   exon            3352..3443
FT   /*tag= e
FT   /label= Exon 1.
FT   /note= "Transit peptide coding region."
FT   intron          3444..3545
FT   /*tag= f
FT   /label= Intron 1.
FT   exon            3546..3608
FT   /*tag= g
FT   /label= Exon 2.
FT   /note= "Transit peptide coding region."
FT   intron          3609..5820
FT   /*tag= h
FT   /label= Intron 2.
FT   exon            5821..6028
FT   /*tag= i
FT   /label= Exon 3.
FT   /note= "Bases 5821-5853 encode the transit peptide,
FT   bases 5854-6028 encode a region of the mature
FT   protein."
FT   intron          6029..6143
FT   /*tag= j
FT   /label= Intron 3
FT   exon            6144..6213
FT   /*tag= k
FT   /label= Exon 4.
FT   /note= "Mature protein coding region."
FT   intron          6214..6647
FT   /*tag= l
FT   /label= Intron 4.
FT   exon            6648..6917
```

```
FT   /*tag=  m
FT   /label= Exon 5.
FT   /note= "Mature protein coding region."
FT   intron          6918..7025
FT   /*tag=  n
FT   /label= Intron 5.
FT   exon            7026..7932
FT   /*tag=  o
FT   /label= Exon 6
FT   /note= "Mature protein coding region."
FT   intron          7933..8244
FT   /*tag=  p
FT   /label= Intron 6.
FT   exon            8245..8361
FT   /*tag=  q
FT   /label= Exon 7.
FT   /note= "Mature protein coding region."
FT   intron          8362..8518
FT   /*tag=  r
FT   /label= Intron 7.
FT   exon            8519..8581
FT   /*tag=  s
FT   /label= Exon 8.
FT   /note= "Mature protein coding region."
FT   intron          8582..9018
FT   /*tag=  t
FT   /label= Intron 8.
FT   exon            9019..9126
FT   /*tag=  u
FT   /label= Exon 9.
FT   /note= "Mature protein coding region."
FT   intron          9127..9594
FT   /*tag=  v
FT   /label= Intron 9.
FT   exon            9595..9696
FT   /*tag=  w
FT   /label= Exon 10
FT   /note= "Mature protein coding region."
FT   intron          9697..9861
FT   /*tag=  x
FT   /label= Intron 10.
FT   exon            9862..9929
FT   /*tag=  y
FT   /label= Exon 11.
FT   /note= "Mature protein coding region."
FT   intron          9930..10010
FT   /*tag=  z
FT   /label= Intron 11.
FT   exon            10011..10091
FT   /*tag=  aa
FT   /label= Exon 12.
FT   /note= "Mature protein coding region."
FT   intron          10092..10209
FT   /*tag=  ab
FT   /label= Intron 12.
FT   exon            10210..10326
FT   /*tag=  ac
FT   /label= Exon 13.
FT   /note= "Mature protein coding region."
FT   intron          10327..10408
FT   /*tag=  ad
FT   /label= Intron 13.
FT   exon            10409..10865
FT   /*tag=  ae
```



```
FT   /label= Exon 14
FT   /note= "Bases 10409-10609 encode a region of the
FT   mature protein.  Bases 10610-10612 are the
FT   translation termination signal, i.e. a stop
FT   codon.
FT   3'UTR           10610..10865
FT   /*tag=  af
FT   polyA_signal    10814..10817
FT   /*tag=  ag
FT   polyA_signal    10839..10842
FT   /*tag=  ah
FT   polyA_signal    10865..10888
FT   /*tag=  ai
PN   J06098656-A.
PD   12-APR-1994.
PF   30-MAR-1992; 102500.
PR   30-MAR-1992; JP-102500.
PA   (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
DR   WPI; 94-155835/19.
PT   Transgenic rice containing the rice starch branch family enzyme -
PT   used to increase the amylopectin content of albumen
PS   Claim 1; Page 16-21; 24pp; Japanese.
CC   The introduction of the rice starch branch-forming enzyme gene into
CC   a rice increase the activity of this enzyme in the plant, thereby
CC   increasing the content of amylopectin in albumen starch and thus
CC   enabling efficient mass production of various proteins.
SQ   Sequence   12151 BP;    3269 A;    2470 C;    2518 G;    3891 T;

  Query Match           53.8%;  Score 14;  DB 12;  Length 12151;
  Best Local Similarity  60.0%;  Pred. No. 3.13e+01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db   9848 aagttccttttcaggttattgtctttgaac 9877
          |||||||| |||        || |  ||||
Qy      2 aagttcctattcNNNNNNNNgtataggaac 31


Search completed: Tue May 14 13:59:39 1996
Job time : 56 secs.
```

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue May 14 13:59:57 1996; MasPar time 25.23 Seconds
484.494 Million cell updates/sec

Tabular output not generated.

Title: >FLP
Description: (1-34) from frt.seq
Perfect Score: 26
N.A. Sequence: 1 gaagttcctattcNNNNNNNNgtataggaacttc 34
Comp: cttcaaggataagNNNNNNNNcatatccttgaag

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 518261 seqs, 179750453 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:STS1 90:STS2 91:STS3 92:STS4
93:STS5 94:STS6
Database: EST-STS-TWO
95:gnEST1 96:gnEST2 97:gnEST3 98:gnEST4 99:gnEST5
100:gnEST6 101:gnEST7 102:gnEST8 103:gnEST9 104:gnEST10
105:gnEST11 106:gnSTS1 107:gnSTS2 108:gnSTS3 109:enEST1
110:enEST2 111:enEST3 112:enEST4 113:enEST5 114:enEST6
115:enEST7 116:enEST8 117:enEST9 118:enEST10 119:enEST11
120:enEST12 121:enEST13 122:enEST14 123:enEST15
124:enEST16 125:enEST17 126:enEST18 127:enEST19

128:enEST20 129:enEST21 130:enSTS1 131:enSTS2 132:enSTS3

Statistics: Mean 6.755; Variance 1.297; scale 5.207

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
c 1	17	65.4	308	64	RICS2602A	Rice cDNA, partial se	3.79e-05
2	16	61.5	438	39	R21554	yg07a01.r1 Homo sapie	9.06e-04
3	16	61.5	462	8	H14983	ym19h09.r1 Homo sapie	9.06e-04
4	16	61.5	463	86	T89822	ye11e02.r1 Homo sapie	9.06e-04
c 5	15	57.7	249	63	RICR2068A	Rice cDNA, partial se	1.91e-02
6	15	57.7	288	35	R07677	ye98d03.r1 Homo sapie	1.91e-02
c 7	15	57.7	288	35	R07677	ye98d03.r1 Homo sapie	1.91e-02
c 8	15	57.7	353	16	H41218	yp64e12.s1 Homo sapie	1.91e-02
c 9	15	57.7	381	119	HS41410	EST79508 Homo sapiens	1.91e-02
c 10	15	57.7	381	70	T29414	EST79508 Homo sapiens	1.91e-02
11	15	57.7	410	56	R79758	yi89e12.r1 Homo sapie	1.91e-02
c 12	15	57.7	412	101	H76310	18015 Arabidopsis tha	1.91e-02
c 13	15	57.7	412	109	AT31015	18015 Arabidopsis tha	1.91e-02
14	15	57.7	448	85	T86566	yd77g07.r1 Homo sapie	1.91e-02
15	15	57.7	452	74	T43599	6862 Arabidopsis thal	1.91e-02
16	15	57.7	452	110	AT5996	6862 Arabidopsis thal	1.91e-02
c 17	15	57.7	469	48	R53335	yg83b07.r1 Homo sapie	1.91e-02
c 18	15	57.7	487	98	H66258	yu18g03.r1 Homo sapie	1.91e-02
c 19	15	57.7	487	116	HS258215	yu18g03.r1 Homo sapie	1.91e-02
20	14	53.8	138	32	HUMGS04157	Human colon 3'directe	3.47e-01
21	14	53.8	138	127	HSGS04157	Human colon 3'directe	3.47e-01
c 22	14	53.8	138	32	HUMGS04157	Human colon 3'directe	3.47e-01
23	14	53.8	162	77	T55306	yb47e03.s1 Homo sapie	3.47e-01
c 24	14	53.8	226	6	H07743	khkh116 Brassica napu	3.47e-01
25	14	53.8	240	114	HS142226	yu86e09.s1 Homo sapie	3.47e-01
26	14	53.8	265	102	H79813	yu10e03.s1 Homo sapie	3.47e-01
27	14	53.8	300	66	T10038	seq1023 Homo sapiens	3.47e-01
28	14	53.8	306	25	HSC1SD111	H. sapiens partial cD	3.47e-01
29	14	53.8	309	88	T98116	ye30b04.r1 Homo sapie	3.47e-01
30	14	53.8	326	108	G11665	human STS WI-10042.	3.47e-01
31	14	53.8	345	73	T40922	ya14c01.s1 Homo sapie	3.47e-01
c 32	14	53.8	347	1	ATTS3932	A. thaliana transcrib	3.47e-01
33	14	53.8	350	27	HSC30G121	H. sapiens partial cD	3.47e-01
c 34	14	53.8	369	114	HS113E01A	Human fetal brain cDN	3.47e-01
35	14	53.8	375	53	R69283	yi39a11.s1 Homo sapie	3.47e-01
c 36	14	53.8	381	41	R27838	yh65h04.s1 Homo sapie	3.47e-01
c 37	14	53.8	390	36	R11119	yf39g03.r1 Homo sapie	3.47e-01
38	14	53.8	418	41	R28016	yh58f05.s1 Homo sapie	3.47e-01
39	14	53.8	421	34	R01221	ye81a01.s1 Homo sapie	3.47e-01
40	14	53.8	450	81	T71405	yd35c09.r1 Homo sapie	3.47e-01
41	14	53.8	468	17	H44838	yp24g12.s1 Homo sapie	3.47e-01
c 42	14	53.8	471	76	T51605	yb27g03.s1 Homo sapie	3.47e-01
43	14	53.8	473	16	H42585	yo09d08.r1 Homo sapie	3.47e-01
c 44	14	53.8	476	4	H01413	yi99c10.r1 Homo sapie	3.47e-01
c 45	14	53.8	516	78	T59688	yc13f05.s1 Homo sapie	3.47e-01

ALIGNMENTS

RESULT 1
LOCUS RICS2602A 308 bp mRNA EST 11-NOV-1994

```
DEFINITION  Rice cDNA, partial sequence (S2602_1A).
ACCESSION   D40544
KEYWORDS    EST(expressed sequence tag).
SOURCE      Oryza sativa (strain Nipponbare) Etiolated shoot (8 days old) cDNA
            to mRNA.
  ORGANISM  Oryza sativa
            Eukaryota; Plantae; Embryobionta; Magnoliophyta; Liliopsida;
            Commelinidae; Cyperales; Poaceae.
REFERENCE   1  (bases 1 to 308)
  AUTHORS   Sasaki,T., Miyao,A. and Yamamoto,K.
  TITLE     Rice cDNA from shoot
  JOURNAL   Unpublished (1994)
COMMENT     PROJECT = RGP
            Submitted (28-OCT-1994) to DDBJ by:
            Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program
            2-1-2, Kannondai,
            Tsukuba, Ibaraki, 305
            Japan
            Phone: 0298-38-7441
            Fax  : 0298-38-7468.

            NCBI gi: 569695
FEATURES             Location/Qualifiers
     source          1..308
                     /organism="Oryza sativa"
                     /strain="Nipponbare"
                     /dev_stage="Etiolated shoot (8 days old)"
                     /sequenced_mol="cDNA to mRNA"
BASE COUNT       95 a     55 c     67 g     88 t      3 others
ORIGIN

  Query Match           65.4%;  Score 17;  DB 64;  Length 308;
  Best Local Similarity  64.5%;  Pred. No. 3.79e-05;
  Matches    20;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db     62 gttccattacagggatatgaagaggaacttc 92
          |||||  |||        ||| |||||||||
Cp     31 gttcctatacNNNNNNNNgaataggaacttc 1


RESULT    2
LOCUS       R21554        438 bp    mRNA            EST       18-APR-1995
DEFINITION  yg07a01.r1 Homo sapiens cDNA clone 31278 5'.
ACCESSION   R21554
KEYWORDS    EST.
SOURCE      human clone=31278 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
```



```
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            GDB: G00-403-625
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 374
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 776335
FEATURES             Location/Qualifiers
     source          1..438
                     /organism="Homo sapiens"
                     /clone="31278"
                     /note="human"
BASE COUNT       99 a     126 c     93 g     119 t      1 others
ORIGIN

  Query Match           61.5%;  Score 16;  DB 39;  Length 438;
  Best Local Similarity  66.7%;  Pred. No. 9.06e-04;
  Matches    16;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     32 ttcctattccctgcaatgtatagg 55
          |||||||||        |||||||
Qy      5 ttcctattcNNNNNNNNgtatagg 28


RESULT    3
LOCUS       H14983        462 bp    mRNA            EST       27-JUN-1995
DEFINITION  ym19h09.r1 Homo sapiens cDNA clone 48660 5'.
ACCESSION   H14983
KEYWORDS    EST.
SOURCE      human clone=48660 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
```

```
COMMENT
            GDB: G00-421-201
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 353
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 879803
FEATURES              Location/Qualifiers
     source           1..462
                      /organism="Homo sapiens"
                      /clone="48660"
                      /note="human"
BASE COUNT      104 a     131 c     100 g     126 t       1 others
ORIGIN

  Query Match           61.5%;  Score 16;  DB 8;  Length 462;
  Best Local Similarity  66.7%;  Pred. No. 9.06e-04;
  Matches    16;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     31 ttcctattccctgcaatgtatagg 54
          |||||||||        |||||||
Qy      5 ttcctattcNNNNNNNNgtatagg 28


RESULT    4
LOCUS       T89822        463 bp    mRNA            EST       20-MAR-1995
DEFINITION  ye11e02.r1 Homo sapiens cDNA clone 117434 5'.
ACCESSION   T89822
KEYWORDS    EST.
SOURCE      human clone=117434 library=Stratagene lung (#937210)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=M13RP1 Rsitel=EcoRI Rsite2=XhoI Normal lung tissue from a 72
            year old male. Cloned unidirectionally. Primer: Oligo dT. Average
            insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
            5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 463)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 252
            Source: IMAGE Consortium, LLNL
```



```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 718335
FEATURES              Location/Qualifiers
     source           1..463
                      /organism="Homo sapiens"
                      /clone="117434"
                      /note="human"
BASE COUNT      105 a     108 c     100 g     143 t       7 others
ORIGIN

  Query Match           61.5%;  Score 16;  DB 86;  Length 463;
  Best Local Similarity  63.3%;  Pred. No. 9.06e-04;
  Matches    19;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db    306 ttccttttcacagggctgtttaggcacttc 335
          ||||| |||        || |||| |||||
Qy      5 ttcctattcNNNNNNNNgtataggaacttc 34


RESULT    5
LOCUS       RICR2068A     249 bp    mRNA            EST       26-MAY-1995
DEFINITION  Rice cDNA, partial sequence (R2068_1A).
ACCESSION   D24501
KEYWORDS    EST(expressed sequence tag).
SOURCE      Oryza sativa (strain Nipponbare, ) Seedling Root cDNA to mRNA.
  ORGANISM  Oryza sativa
            Eukaryotae; mitochondrial eukaryotes; Chlorophyta/Embryophyta
            group; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Commelinidae; Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 249)
  AUTHORS   Minobe,Y. and Sasaki,T.
  TITLE     Rice cDNA from root
  JOURNAL   Unpublished (1993)
COMMENT     Submitted (2-NOV-1993) to DDBJ by:
            Yuzo Minobe
            Dept. Rice Genome Research Program
            National Institute of Agrobiological Resources
            Kannondai 2-1-2
            Tsukuba, Ibaraki
            Japan
            Phone: 0298-38-7441
            Fax:   0298-38-7468
            PROJECT ='RGP'.

            NCBI gi: 428353
FEATURES              Location/Qualifiers
     source           1..249
                      /organism="Oryza sativa"
                      /strain="Nipponbare"
                      /dev_stage="Seedling"
                      /sequenced_mol="cDNA to mRNA"
                      /tissue_type="Root"
BASE COUNT       86 a      40 c      60 g      61 t       2 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 63;  Length 249;
  Best Local Similarity  60.6%;  Pred. No. 1.91e-02;
  Matches    20;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db    145 aagttactataatggttggcaaagaggaaattc 177
          ||||| |||||          || ||||| |||
Cp     33 aagttcctatacNNNNNNNNgaataggaacttc 1
```

```
RESULT    6
LOCUS       R07677        288 bp    mRNA            EST       05-APR-1995
DEFINITION  ye98d03.r1 Homo sapiens cDNA clone 125765 5'.
ACCESSION   R07677
KEYWORDS    EST.
SOURCE      human clone=125765 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 288)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 248
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 759600
FEATURES             Location/Qualifiers
     source          1..288
                     /organism="Homo sapiens"
                     /clone="125765"
                     /note="human"
BASE COUNT       90 a     49 c     45 g    104 t
ORIGIN

  Query Match           57.7%;  Score 15;  DB 35;  Length 288;
  Best Local Similarity  60.6%;  Pred. No. 1.91e-02;
  Matches    20;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db     49 gaagttcctattacttcttatgattgggagctt 81
          ||||||||||||         |  | ||| |||
Qy      1 gaagttcctattcNNNNNNNNgtataggaactt 33


RESULT    7
LOCUS       R07677        288 bp    mRNA            EST       05-APR-1995
DEFINITION  ye98d03.r1 Homo sapiens cDNA clone 125765 5'.
ACCESSION   R07677
KEYWORDS    EST.
```



```
SOURCE      human clone=125765 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 288)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 248
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 759600
FEATURES             Location/Qualifiers
     source          1..288
                     /organism="Homo sapiens"
                     /clone="125765"
                     /note="human"
BASE COUNT       90 a     49 c     45 g    104 t
ORIGIN

  Query Match           57.7%;  Score 15;  DB 35;  Length 288;
  Best Local Similarity  60.6%;  Pred. No. 1.91e-02;
  Matches    20;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db     49 gaagttcctattacttcttatgattgggagctt 81
          ||||||||||||         |  | ||| |||
Cp     34 gaagttcctatacNNNNNNNNgaataggaactt 2


RESULT    8
LOCUS       H41218        353 bp    mRNA            EST       16-AUG-1995
DEFINITION  yp64e12.s1 Homo sapiens cDNA clone 192238 3' similar to contains
            Alu repetitive element;.
ACCESSION   H41218
KEYWORDS    EST.
SOURCE      human clone=192238 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=Promega -21m13 Rsite1=Pac I
            Rsite2=Eco RI Liver and spleen from a 20 week-post conception male
            fetus. 1st strand cDNA was primed with a Pac I - oligo(dT) primer
            [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
```

```
            double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
            digested with Pac I and cloned into the Pac I and Eco RI sites of
            the modified pT7T3 vector.  Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 353)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 58
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 917270
FEATURES             Location/Qualifiers
     source          1..353
                     /organism="Homo sapiens"
                     /clone="192238"
                     /note="human"
BASE COUNT       91 a     68 c     74 g    106 t     14 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 16;  Length 353;
  Best Local Similarity  64.3%;  Pred. No. 1.91e-02;
  Matches    18;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    224 gaagatccnatacagngataaaaatagg 251
          |||| ||| ||||  |      ||||||
Cp     34 gaagttcctatacNNNNNNNNgaatagg 7


RESULT    9
ID   HS41410    standard; RNA; EST; 381 BP.
AC   T29414;
DT   09-JAN-1995 (Rel. 42, Created)
DT   08-SEP-1995 (Rel. 45, Last updated, Version 2)
DE   EST79508 Homo sapiens cDNA similar to beta-1-glycoprotein 4,
DE   pregnancy-specific (HT:2583).
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC   Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
RN   [1]
RP   1-381
RA   Adams M.D., Kerlavage A.R., Fleischmann R.D., Fuldner R.A.,
RA   Bult C.J., Lee N., Kirkness E.F., Weinstock K.G., Gocayne J.D.,
```



```
RA   White O., Sutton G., Blake J.A., Brandon R.C., Chiu M.W.,
RA   Clayton R.A., Cline R.T., Cotton M.D., Earle-Hughes J., Fine L.D.,
RA   FitzGerald L.M., FitzHugh W.M., Fritchman J.L., Geoghagen N.S.M.,
RA   Glodek A., Gnehm C.L., Hanna M.C., Hedblom E., Hinkle Jr P.S.,
RA   Kelley J.M., Klimek K.M., Kelley J.C., Liu L.I., Marmaros S.M.,
RA   Merrick J.M., MORENO-PALANQUES R.F., McDonald L.A., Nguyen D.T.,
RA   Pellegrino S.M., Phillips C.A., Ryder S.E., Scott J.L.,
RA   Saudek D.M., Shirley R., Small K.V., Spriggs T.A., Utterback T.R.,
RA   Weidman J.F., Li Y., Bednarik D.P., Cao L., Cepeda M.A.,
RA   Coleman T.A., Collins E.J., Dimke D., Feng P., Ferrie A.,
RA   Fischer C., Hastings G.A., He W.W., Hu J.S., Greene J.M.,
RA   Gruber J., Hudson P., Kim A., Kozak D.L., Kunsch C., Ji H., Li H.,
RA   Meissner P.S., Olsen H., Raymond L., Wei Y.F., Wing J., Xu C.,
RA   Yu G.L., Ruben S.M., Dillon P.J., Fannon M.R., Rosen C.A.,
RA   Haseltine W.A., Fields C., Fraser C.M., Venter J.C.;
RT   "Initial Assessment of Human Gene Diversity and Expression
RT   Patterns Based Upon 52 Million Basepairs of cDNA Sequence";
RL   Unpublished.
CC   Other_ESTs: THC24005 Contact: Venter, JC The Institute for Genomic
CC   Research 932 Clopper Rd, Gaithersburg, MD 20878 Tel: 3018699056
CC   Fax: 3018699423 Email: tdbinfo@tdb.tigr.org For clone availability,
CC   additional sequence and expression information related to this EST,
CC   please contact the TIGR Database (tdbinfo@tdb.tigr.org). NCBI gi:
CC   611512
FH   Key             Location/Qualifiers
FH
FT   source          1..381
FT                   /organism="Homo sapiens"
FT                   /note="human"
FT   mRNA            <1..>381
SQ   Sequence 381 BP; 107 A; 78 C; 71 G; 125 T; 0 other;

  Query Match           57.7%;  Score 15;  DB 119;  Length 381;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    313 gaagctcctatactacatgtggaat 337
          |||| ||||||||        ||||
Cp     34 gaagttcctatacNNNNNNNNgaat 10


RESULT   10
LOCUS       T29414         381 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST79508 Homo sapiens cDNA similar to beta-1-glycoprotein 4,
            pregnancy-specific (HT:2583).
ACCESSION   T29414
KEYWORDS    EST.
SOURCE      human primer=M13-21 library=Human Placenta.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 381)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
```

```
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: THC24005
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org).

            NCBI gi: 611512
FEATURES             Location/Qualifiers
     source          1..381
                     /organism="Homo sapiens"
                     /note="human"
     mRNA            <1..>381
BASE COUNT      107 a      78 c      71 g     125 t
ORIGIN

  Query Match           57.7%;  Score 15;  DB 70;  Length 381;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    313 gaagctcctatactacatgtggaat 337
          |||| ||||||||        ||||
Cp     34 gaagttcctatacNNNNNNNNgaat 10


RESULT   11
LOCUS       R79758        410 bp    mRNA            EST       09-JUN-1995
DEFINITION  yi89e12.r1 Homo sapiens cDNA clone 146446 5'.
ACCESSION   R79758
KEYWORDS    EST.
SOURCE      human clone=146446 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsitel=Not I Rsite2=Eco RI Female placenta
            obtained at birth (full term). 1st strand cDNA was primed with a
            Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 410)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
```



```
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 277
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 856039
FEATURES             Location/Qualifiers
     source          1..410
                     /organism="Homo sapiens"
                     /clone="146446"
                     /note="human"
BASE COUNT      117 a      82 c      90 g     118 t      3 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 56;  Length 410;
  Best Local Similarity  63.0%;  Pred. No. 1.91e-02;
  Matches    17;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    232 tccacttcgtacttctgtataggaact 258
          ||| |||         ||||||||||
Qy      6 tcctattcNNNNNNNNgtataggaact 32


RESULT   12
LOCUS       H76310        412 bp    mRNA            EST       06-NOV-1995
DEFINITION  18015 Arabidopsis thaliana cDNA clone 200C2T7.
ACCESSION   H76310
KEYWORDS    EST.
SOURCE      thale cress clone=200C2T7 primer=T7 dye primer library=Lambda-PRL2
            strain=var columbia vector=lambda Zip-Lox Rsitel=Sal Rsite2=Not
            Lambda PRL2 is a cDNA library derived from equal quantities of 4
            pools of mRNA.  The mRNA sources were 1) 7 day germinated etiolated
            seedlings; 2) tissue culture grown roots; 3) staged plants half
            with 24 hour light cycle, half on 16 hr light, 8 hour dark-
            rosettes;  4) same plants as 3 but aerial tissue (stems, flowers
            and siliques.  The vector is BRL's lambda Zip-Lox.  The cDNA
            inserts were directionally cloned with Sal-Not arms using oligo dT
            primed cDNA.
  ORGANISM  Arabidopsis thaliana
            Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Capparales;
            Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Newman,T., de Bruijn,F.J., Green,P., Keegstra,K., Kende,H.,
            McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
            Retzel,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
COMMENT
            Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
```

```
           Lansing,Mi
           Tel: 517-353-0854
           Fax: 517-353-9168
           Email: 22313tcn@ibm.cl.msu.edu.

           NCBI gi: 1053561
FEATURES             Location/Qualifiers
     source          1..412
                     /organism="Arabidopsis thaliana"
                     /clone="200C2T7"
                     /strain="var columbia"
                     /note="thale cress"
     mRNA            <1..>412
BASE COUNT      103 a     83 c     84 g    126 t     16 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 101;  Length 412;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    190 atacattacaaagaacaggaacttc 214
          ||||        ||| |||||||||
Cp     25 atacNNNNNNNNgaataggaacttc 1


RESULT   13
ID   AT31015    standard; RNA; EST; 412 BP.
AC   H76310;
DT   10-NOV-1995 (Rel. 45, Created)
DT   10-NOV-1995 (Rel. 45, Last updated, Version 1)
DE   18015 Arabidopsis thaliana cDNA clone 200C2T7.
KW   EST.
OS   Arabidopsis thaliana
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
RN   [1]
RP   1-412
RA   Newman T., de Bruijn F.J., Green P., Keegstra K., Kende H.,
RA   McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomashow M.,
RA   Retzel E., Somerville C.;
RT   "Genes galore: a summary of methods for accessing results from
RT   large-scale partial sequencing of anonymous Arabidopsis cDNA
RT   clones";
RL   Plant Physiol. 106:1241-1255(1994).
CC   Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan
CC   State University MSU-DOE-PRL, Michigan State University,Plant
CC   Biology Bldg.,E. Lansing,Mi Tel: 517-353-0854 Fax: 517-353-9168
CC   Email: 22313tcn@ibm.cl.msu.edu. NCBI gi: 1053561
FH   Key             Location/Qualifiers
FH
FT   source          1..412
FT                   /organism="Arabidopsis thaliana"
FT                   /clone="200C2T7"
FT                   /strain="var columbia"
FT                   /note="thale cress"
FT   mRNA            <1..>412
SQ   Sequence 412 BP; 103 A; 83 C; 84 G; 126 T; 16 other;

  Query Match           57.7%;  Score 15;  DB 109;  Length 412;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    190 atacattacaaagaacaggaacttc 214
          ||||        ||| |||||||||
```



```
Cp     25 atacNNNNNNNNgaataggaacttc 1


RESULT   14
LOCUS       T86566        448 bp    mRNA            EST       17-MAR-1995
DEFINITION  yd77g07.r1 Homo sapiens cDNA clone 114300 5' similar to
            SP:UBC5_DROME P35128 UBIQUITIN-CONJUGATING ENZYME E2-17 KD
            ;contains Alu repetitive element;.
ACCESSION   T86566
KEYWORDS    EST.
SOURCE      human clone=114300 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 426
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

           NCBI gi: 714918
FEATURES             Location/Qualifiers
     source          1..448
                     /organism="Homo sapiens"
                     /clone="114300"
                     /note="human"
BASE COUNT      110 a     123 c     97 g    114 t      4 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 85;  Length 448;
  Best Local Similarity  62.1%;  Pred. No. 1.91e-02;
  Matches    18;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db    414 gttccttttgtgatttctgtatagggact 442
          |||||| ||         ||||||| |||
Qy      4 gttcctattcNNNNNNNNgtataggaact 32


RESULT   15
LOCUS       T43599        452 bp    mRNA            EST       17-AUG-1995
```

```
DEFINITION  6862 Arabidopsis thaliana cDNA clone 121G9T7.
ACCESSION   T43599
KEYWORDS    EST.
SOURCE      thale cress clone=121G9T7 library=Lambda-PRL2 strain=var columbia
            vector=lambda Zip-Lox primer=T7 dye primer Rsite1=Sal Rsite2=Not
            Lambda PRL2 is a cDNA library derived from equal quantities of 4
            pools of mRNA.  The mRNA sources were 1) 7 day germinated etiolated
            seedlings; 2) tissue culture grown roots; 3) staged plants half
            with 24 hour light cycle, half on 16 hr light, 8 hour dark-
            rosettes;  4) same plants as 3 but aerial tissue (stems, flowers
            and siliques.  The vector is BRL's lambda Zip-Lox.  The cDNA
            inserts were directionally cloned with Sal-Not arms using oligo dT
            primed cDNA.
  ORGANISM  Arabidopsis thaliana
            Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Capparales;
            Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   Newman,T., de Bruijn,F.J., Green,P., Keegstra,K., Kende,H.,
            McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
            Retzel,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
COMMENT
            Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
            Lansing,Mi
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcn@ibm.cl.msu.edu.

            NCBI gi: 947993
FEATURES              Location/Qualifiers
     source           1..452
                      /organism="Arabidopsis thaliana"
                      /clone="121G9T7"
                      /strain="var columbia"
                      /note="thale cress"
BASE COUNT      116 a     69 c    115 g    130 t     22 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 74;  Length 452;
  Best Local Similarity  64.5%;  Pred. No. 1.91e-02;
  Matches    20;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db     19 agttgctattctntctactggagatgaactt 49
          |||| |||||| |      | | | ||||||
Qy      3 agttcctattcNNNNNNNNgtataggaactt 33


RESULT   16
ID   AT5996     standard; RNA; EST; 452 BP.
AC   T43599;
DT   03-FEB-1995 (Rel. 42, Created)
DT   10-NOV-1995 (Rel. 45, Last updated, Version 8)
DE   6862 Arabidopsis thaliana cDNA clone 121G9T7.
KW   EST.
OS   Arabidopsis thaliana
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
RN   [1]
RP   1-452
```



```
RA   Newman T., de Bruijn F.J., Green P., Keegstra K., Kende H.,
RA   McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomashow M.,
RA   Retzel E., Somerville C.;
RT   "Genes galore: a summary of methods for accessing results from
RT   large-scale partial sequencing of anonymous Arabidopsis cDNA
RT   clones";
RL   Plant Physiol. 106:1241-1255(1994).
DR   AGIS; T43599; AGIS July 1995.
CC   Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan
CC   State University MSU-DOE-PRL, Michigan State University,Plant
CC   Biology Bldg.,E. Lansing,Mi Tel: 517-353-0854 Fax: 517-353-9168
CC   Email: 22313tcn@ibm.cl.msu.edu. NCBI gi: 947993
FH   Key             Location/Qualifiers
FH
FT   source          1..452
FT                   /organism="Arabidopsis thaliana"
FT                   /clone="121G9T7"
FT                   /strain="var columbia"
FT                   /note="thale cress"
SQ   Sequence 452 BP; 116 A; 69 C; 115 G; 130 T; 22 other;

  Query Match           57.7%;  Score 15;  DB 110;  Length 452;
  Best Local Similarity  64.5%;  Pred. No. 1.91e-02;
  Matches    20;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db     19 agttgctattctntctactggagatgaactt 49
          |||| |||||| |      | | | ||||||
Qy      3 agttcctattcNNNNNNNNgtataggaactt 33


RESULT   17
LOCUS       R53335       469 bp    mRNA            EST       18-MAY-1995
DEFINITION  yg83b07.r1 Homo sapiens cDNA clone 39926 5'.
ACCESSION   R53335
KEYWORDS    EST.
SOURCE      human clone=39926 library=Soares infant brain 1NIB vector=Lafmid BA
            host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
            Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded
            cDNA was ligated to Hind III adaptors (Pharmacia), digested with
            Not I and directionally cloned into the Not I and Hind III sites of
            the Lafmid BA vector. Library went through one round of
            normalization. Library constructed by Bento Soares and M.Fatima
            Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT

            GDB: GOO-412-467
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
```

```
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 359
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 815237
FEATURES              Location/Qualifiers
     source           1..469
                      /organism="Homo sapiens"
                      /clone="39926"
                      /note="human"
BASE COUNT      124 a    102 c    107 g    133 t      3 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 48;  Length 469;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    306 ataataaccagagaataggaacttc 330
          |||         |||||||||||||
Cp     25 atacNNNNNNNNgaataggaacttc 1


RESULT   18
LOCUS       H66258        487 bp    mRNA            EST       18-OCT-1995
DEFINITION  yu18g03.r1 Homo sapiens cDNA clone 234196 5' similar to gb:S52450
            NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (HUMAN);.
ACCESSION   H66258
KEYWORDS    EST.
SOURCE      human clone=234196 primer=M13RP1 library=Soares fetal liver spleen
            1NFLS vector=pT7T3D (Pharmacia) with a modified polylinker
            host=DH10B (ampicillin resistant) Rsite1=Pac I Rsite2=Eco RI Liver
            and spleen from a 20 week-post conception male fetus. 1st strand
            cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 379
```



```
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 1024998
FEATURES              Location/Qualifiers
     source           1..487
                      /organism="Homo sapiens"
                      /clone="234196"
                      /note="human"
     mRNA             <1..>487
BASE COUNT      140 a     93 c    109 g    144 t      1 others
ORIGIN

  Query Match           57.7%;  Score 15;  DB 98;  Length 487;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    394 ttcctatgcttctgggtgaatagga 418
          ||||||| |        ||||||||
Cp     30 ttcctatacNNNNNNNNgaatagga 6


RESULT   19
ID   HS258215   standard; RNA; EST; 487 BP.
AC   H66258;
DT   21-OCT-1995 (Rel. 45, Created)
DT   21-OCT-1995 (Rel. 45, Last updated, Version 1)
DE   yu18g03.r1 Homo sapiens cDNA clone 234196 5' similar to gb:S52450
DE   NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (HUMAN);.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC   Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
RN   [1]
RP   1-487
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA   Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA   Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA   Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project";
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu High quality sequence stops: 379 Source: IMAGE
CC   Consortium, LLNL This clone is available royalty-free through LLNL
CC   ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC   information. NCBI gi: 1024998
FH   Key             Location/Qualifiers
FH
FT   source          1..487
FT                   /organism="Homo sapiens"
FT                   /clone="234196"
FT                   /note="human"
FT   mRNA            <1..>487
SQ   Sequence 487 BP; 140 A; 93 C; 109 G; 144 T; 1 other;

  Query Match           57.7%;  Score 15;  DB 116;  Length 487;
  Best Local Similarity  64.0%;  Pred. No. 1.91e-02;
  Matches    16;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    394 ttcctatgcttctgggtgaatagga 418
          ||||||| |        ||||||||
```

```
Cp     30 ttcctatacNNNNNNNNgaatagga 6


RESULT   20
LOCUS       HUMGS04157    138 bp    mRNA            EST       18-JUN-1994
DEFINITION  Human colon 3'directed MboI cDNA, HUMGS04157, clone cm1934.
ACCESSION   D25789
KEYWORDS    EST(expressed sequence tag); colon; endothel; gene signature(GS).
SOURCE      Homo sapiens male adult colon mucosa cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 138)
  AUTHORS   Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.
  TITLE     Global analysis of gene expression in colon mucosa: a large scale
            random cDNA sequencing analysis
  JOURNAL   Unpublished (1993)
COMMENT     Submitted (22-Nov-1993) to DDBJ by: Kousaku Okubo
            Institute for Molecular and Cellular Biology
            Osaka University
            3-1, Yamadaoka
            Suita, Osaka, 565
            Japan
            Phone: 06-877-5111
            Fax  : 06-875-1922.

            NCBI gi: 500472
FEATURES             Location/Qualifiers
     source          1..138
                     /organism="Homo sapiens"
                     /dev_stage="adult"
                     /sequenced_mol="cDNA to mRNA"
                     /sex="male"
                     /tissue_type="colon mucosa"
BASE COUNT       45 a     23 c     28 g     36 t      6 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 32;  Length 138;
  Best Local Similarity  58.1%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db     98 agttctnattcatttgaataaaaaggaactt 128
          |||||  ||||          | ||||||||
Qy      3 agttcctattcNNNNNNNNgtataggaactt 33


RESULT   21
ID   HSGS04157  standard; RNA; EST; 138 BP.
AC   D25789;
DT   23-JUN-1994 (Rel. 40, Created)
DT   27-NOV-1995 (Rel. 45, Last updated, Version 2)
DE   Human colon 3'directed MboI cDNA, HUMGS04157, clone cm1934.
KW   colon; endothel; EST(expressed sequence tag); gene signature(GS).
OS   Homo sapiens (human)
OC   Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC   Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
RN   [1]
RP   1-138
RA   Okubo K., Itoh K., Yoshii J., Yokouchi H., Matsubara K.;
RT   "Global analysis of gene expression in colon mucosa: a large scale
RT   random cDNA sequencing analysis";
RL   Unpublished.
RN   [2]
RA   Okubo K., Yoshii J., Yokouchi H., Kameyama M., Matsubara K.;
```



```
RT   "An expression profile of active genes in human colonic mucosa";
RL   DNA Res. 1:37-45(1994).
CC   Submitted (22-Nov-1993) to DDBJ by: Kousaku Okubo Institute for
CC   Molecular and Cellular Biology Osaka University 3-1, Yamadaoka
CC   Suita, Osaka, 565 Japan Phone: 06-877-5111 Fax : 06-875-1922
FH   Key             Location/Qualifiers
FH
FT   source          1..138
FT                   /organism="Homo sapiens"
FT                   /dev_stage="adult"
FT                   /sequenced_mol="cDNA to mRNA"
FT                   /sex="male"
FT                   /tissue_type="colon mucosa"
SQ   Sequence 138 BP; 45 A; 23 C; 28 G; 36 T; 6 other;

  Query Match           53.8%;  Score 14;  DB 127;  Length 138;
  Best Local Similarity  58.1%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db     98 agttctnattcatttgaataaaaaggaactt 128
          |||||  ||||          | ||||||||
Qy      3 agttcctattcNNNNNNNNgtataggaactt 33


RESULT   22
LOCUS       HUMGS04157    138 bp    mRNA            EST       18-JUN-1994
DEFINITION  Human colon 3'directed MboI cDNA, HUMGS04157, clone cm1934.
ACCESSION   D25789
KEYWORDS    EST(expressed sequence tag); colon; endothel; gene signature(GS).
SOURCE      Homo sapiens male adult colon mucosa cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 138)
  AUTHORS   Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.
  TITLE     Global analysis of gene expression in colon mucosa: a large scale
            random cDNA sequencing analysis
  JOURNAL   Unpublished (1993)
COMMENT     Submitted (22-Nov-1993) to DDBJ by: Kousaku Okubo
            Institute for Molecular and Cellular Biology
            Osaka University
            3-1, Yamadaoka
            Suita, Osaka, 565
            Japan
            Phone: 06-877-5111
            Fax  : 06-875-1922.

            NCBI gi: 500472
FEATURES             Location/Qualifiers
     source          1..138
                     /organism="Homo sapiens"
                     /dev_stage="adult"
                     /sequenced_mol="cDNA to mRNA"
                     /sex="male"
                     /tissue_type="colon mucosa"
BASE COUNT       45 a     23 c     28 g     36 t      6 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 32;  Length 138;
  Best Local Similarity  58.1%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db     98 agttctnattcatttgaataaaaaggaactt 128
          |||||  || |          || ||||||||
```

```
Cp     32 agttcctatacNNNNNNNNgaataggaactt 2


RESULT   23
LOCUS       T55306        162 bp    mRNA            EST       08-FEB-1995
DEFINITION  yb47e03.s1 Homo sapiens cDNA clone 74332 3' similar to gb:X69150
            40S RIBOSOMAL PROTEIN S18 (HUMAN).
ACCESSION   T55306
KEYWORDS    EST.
SOURCE      human clone=74332 library=Stratagene fetal spleen (#937205)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=-21m13 Rsite1=EcoRI Rsite2=XhoI Pooled fetal spleens. Cloned
            unidirectionally.  Primer: Oligo dT. Average insert size: 1.0 kb;
            Uni-ZAP XR Vector; 5' adaptor sequence: 5'-GAATTCGGCACGAG-3'; 3'
            adaptor sequence: 5'-CTCGAGTTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 162)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence starts: 1
            High qality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 657167
FEATURES              Location/Qualifiers
     source           1..162
                      /organism="Homo sapiens"
                      /clone="74332"
                      /note="human"
BASE COUNT       31 a     37 c     32 g     49 t     13 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 77;  Length 162;
  Best Local Similarity  57.7%;  Pred. No. 3.47e-01;
  Matches    15;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db     96 ctnttcaagaaccagtntgggaactt 121
          || |||        || | |||||||
Qy      8 ctattcNNNNNNNNgtataggaactt 33


RESULT   24
LOCUS       H07743        226 bp    mRNA            EST       23-JUN-1995
DEFINITION  khkh116 Brassica napus cDNA 3'.
ACCESSION   H07743
KEYWORDS    EST.
SOURCE      rape library=BNL1 strain=cv. Naehan vector=pT7T3D host=NM522
            primer=M13 forward Rsite1=NotI Rsite2=EcoRI Poly(A)-mRNA  was
```



```
            purified from  the  leaf  of B.napus.  cDNA  library  was
            constructed from  the mRNAs  by oligo(dT)  priming and
            directionally  cloned from the NotI site in the vector pT7T3D
            (Pharmacia) to the EcoRI site.
  ORGANISM  Brassica napus
            Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Dilleniidae;
            Capparales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 226)
  AUTHORS   Sohn,U., Lee,C.M., Cho,K.H., Jeon,Y.H., Hahn,T.R. and Nam,H.G.
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Uik Sohn
            Laboratory of Molecular Biology
            Kyungpook National University
            Dept. of Genetic Eng.,Kyungpook National Univ.,Taegu 702-701,Korea
            Tel: 0539505382
            Fax: 0539555327
            Email: usohn@bh.kyungpook.ac.kr
            EST is putatively homologous to unknown gene.

            NCBI gi: 872565
FEATURES              Location/Qualifiers
     source           1..226
                      /organism="Brassica napus"
                      /strain="cv. Naehan"
                      /note="rape"
BASE COUNT       61 a     45 c     66 g     54 t
ORIGIN

  Query Match           53.8%;  Score 14;  DB 6;  Length 226;
  Best Local Similarity  58.8%;  Pred. No. 3.47e-01;
  Matches    20;  Conservative    0;  Mismatches  14;  Indels   0;  Gaps   0;

Db    130 gaagttcccagacttcagttcgagttagaaccctc 163
          |||||||| | ||        || |  |||| ||
Cp     34 gaagttcctatacNNNNNNNNgaataggaacttc 1


RESULT   25
ID   HS142226   standard; RNA; EST; 240 BP.
AC   H78142;
DT   15-NOV-1995 (Rel. 45, Created)
DT   15-NOV-1995 (Rel. 45, Last updated, Version 1)
DE   yu86e09.s1 Homo sapiens cDNA clone 240712 3' similar to gb:L11244
DE   C4B-BINDING PROTEIN BETA CHAIN PRECURSOR (HUMAN);.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC   Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
RN   [1]
RP   1-240
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA   Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA   Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA   Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project";
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu High quality sequence stops: 59 Source: IMAGE
CC   Consortium, LLNL This clone is available royalty-free through LLNL
CC   ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC   information. NCBI gi: 1056231
```

```
FH   Key             Location/Qualifiers
FH
FT   source          1..240
FT                   /organism="Homo sapiens"
FT                   /clone="240712"
FT                   /note="human"
FT   mRNA            <1..>240
SQ   Sequence 240 BP; 54 A; 58 C; 44 G; 71 T; 13 other;

  Query Match           53.8%;  Score 14;  DB 114;  Length 240;
  Best Local Similarity  61.8%;  Pred. No. 3.47e-01;
  Matches    21;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db    178 gaagtgccttctcanactctggtttgggagcttc 211
          ||||| |||  || |      || | ||| ||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34


RESULT   26
LOCUS       H79813        265 bp    mRNA            EST       09-NOV-1995
DEFINITION  yu10e03.s1 Homo sapiens cDNA clone 233404 3' similar to
            gb|M87942|HUMAALU83 Human carcinoma cell-derived Alu RNA
            transcript, (rRNA); gb:X62078 GANGLIOSIDE GM2 ACTIVATOR PRECURSOR
            (HUMAN);contains Alu repetitive element;.
ACCESSION   H79813
KEYWORDS    EST.
SOURCE      human clone=233404 primer=Promega -21m13 library=Soares fetal liver
            spleen 1NFLS vector=pT7T3D (Pharmacia) with a modified polylinker
            host=DH10B (ampicillin resistant) Rsite1=Pac I Rsite2=Eco RI Liver
            and spleen from a 20 week-post conception male fetus. 1st strand
            cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 231
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 1057902
FEATURES              Location/Qualifiers
```



```
     source           1..265
                      /organism="Homo sapiens"
                      /clone="233404"
                      /note="human"
     mRNA             <1..>265
BASE COUNT       91 a     50 c     57 g     60 t      7 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 102;  Length 265;
  Best Local Similarity  68.2%;  Pred. No. 3.47e-01;
  Matches    15;  Conservative    0;  Mismatches   7;  Indels   0;  Gaps   0;

Db    217 cacagntttgtataggaacttc 238
          |    |   ||||||||||||||
Qy     13 cNNNNNNNNgtataggaacttc 34


RESULT   27
LOCUS       T10038        300 bp    mRNA            EST       13-APR-1994
DEFINITION  seq1023 Homo sapiens cDNA clone b4HB3MA-COT8-HAP-Ft57 3'.
ACCESSION   T10038
KEYWORDS    EST.
SOURCE      human clone=b4HB3MA-COT8-HAP-Ft57 library=b4HB3MA Cot8-HAP-Ft
            vector=Lafmid BA host=E. coli primer=M13 Forward (Universal)
            Rsite1=Not I Rsite2=Hind III Size-selected cDNA from polyA+ RNA
            from human brain. 3-month old neonate.This library is the result of
            an attempted normalization of library b4HB3MA.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Soares,M.B., de Fatima Bonaldo,M., Jelenc,P., Long Su,L.L. and
            Efstratiadis,A.
  TITLE     Construction and characterization of a normalized cDNA library
  JOURNAL   Unpublished (1993)
COMMENT
            Contact: Bento Soares
            Columbia University
            Department of Psychiatry,722 W 168th Street, Unit #41, New York,NY
            10032
            Tel: 2129602313
            Fax: 2127955886
            Email: cuc@cuccfa.ccc.columbia.edu.

            NCBI gi: 471389
FEATURES              Location/Qualifiers
     source           1..300
                      /organism="Homo sapiens"
                      /clone="b4HB3MA-COT8-HAP-Ft57"
                      /note="human"
BASE COUNT      103 a     53 c     51 g     93 t
ORIGIN

  Query Match           53.8%;  Score 14;  DB 66;  Length 300;
  Best Local Similarity  61.5%;  Pred. No. 3.47e-01;
  Matches    16;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    186 cctattcttgcacaggaataggacct 211
          |||||||        | ||||||  ||
Qy      7 cctattcNNNNNNNNgtataggaact 32


RESULT   28
LOCUS       HSC1SD111     306 bp    RNA             EST       15-FEB-1995
```

```
DEFINITION  H. sapiens partial cDNA sequence; clone c-1sd11.
ACCESSION   F07017
KEYWORDS    partial cDNA sequence; transcribed sequence fragment.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 306)
  AUTHORS   Genexpress.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-1995) to the EMBL/GenBank/DDBJ databases.
            Genethon, B.P. 60, 91002 Evry Cedex France and Genetique
            Moleculaire et Biologie du developpement, CNRS UPR420 B.P. 8, 94801
            Villejuif Cedex France. E-mail: genexpress@genethon.fr
REFERENCE   2  (bases 1 to 306)
  AUTHORS   Genexpress.
  TITLE     The Genexpress cDNA program
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 306)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Boucher,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,
            Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: Integated molecular analysis of the human genome and its
            expression
  JOURNAL   C.R. Acad. Sci., III, Sci. Vie 318, 263-272 (1995)
COMMENT     Cloning_method: total mRNA was oligo-(dT) primed and directionally
            cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA
            vector;
            Sequencing_method: single read, full automatic;
            Primer: M13_reverse
            cDNA sequence colinear to mRNA
            Stretch_removed: nothing
            Normalization_method: Bento Soares, P.N.A.S. 91:9228-9232(1994);
            Genexpress_library_idt: C;
            Genexpress_sequence_idt: y1c-1sd11.

            NCBI gi: 672657
FEATURES              Location/Qualifiers
     source           1..306
                      /organism="Homo sapiens"
                      /clone_lib="normalized infant brain cDNA from B.Soares,
                      Psychiatry Dept. Columbia University USA"
                      /sex="female"
                      /tissue_type="total brain"
                      /dev_stage="3 months old"
                      /isolate="muscular atrophy patient"
BASE COUNT       95 a     74 c     71 g     66 t
ORIGIN

  Query Match           53.8%;  Score 14;  DB 25;  Length 306;
  Best Local Similarity  60.0%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    237 gttcccattctgtgtgcagcctagggactt 266
          ||||| ||||        |  |||| ||||
Qy      4 gttcctattcNNNNNNNNgtataggaactt 33


RESULT   29
LOCUS       T98116        309 bp    mRNA            EST       29-MAR-1995
```



```
DEFINITION  ye30b04.r1 Homo sapiens cDNA clone 119215 5' similar to contains L1
            repetitive element ;.
ACCESSION   T98116
KEYWORDS    EST.
SOURCE      human clone=119215 library=Stratagene lung (#937210)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=M13RP1 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72
            year old male. Cloned unidirectionally. Primer: Oligo dT. Average
            insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
            5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 196
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 747461
FEATURES              Location/Qualifiers
     source           1..309
                      /organism="Homo sapiens"
                      /clone="119215"
                      /note="human"
BASE COUNT       88 a     75 c     35 g    110 t      1 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 88;  Length 309;
  Best Local Similarity  59.4%;  Pred. No. 3.47e-01;
  Matches    19;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db    254 aaattcctaatctgttttttctatagggcctt 285
          || |||||| ||         ||||||  |||
Qy      2 aagttcctattcNNNNNNNNgtataggaactt 33


RESULT   30
LOCUS       G11665        326 bp    DNA             STS       19-OCT-1995
DEFINITION  human STS WI-10042.
ACCESSION   G11665
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human STSs derived from random genomic DNA.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 326)
```

```
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped STSs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: GATACTTCCTTTTCATACAAATTGG
            Primer B: CCTCTGCTGAAAGCAAAGTAA
            STS size: 200
            PCR Profile:
                  Presoak:
                  Denaturation:
                  Annealing: 56 degrees C
                  Polymerization:
                  PCR Cycles: 35
                  Thermal Cycler:
            Protocol:
                Template: 10 ng
                Primer: each 5 pM
                dNTPs: each 4 nM
                Taq Polymerase: 0.025 units/ul
                Total Vol: 20 ul

            Buffer:
                MgCl2: 1.5 mM
                KCl: 50 mM
                Tris-HCL: 10 mM
                pH: 9.3

            Prepared with primer pairs derived from random genomic sequence.

            NCBI gi: 1022420
FEATURES              Location/Qualifiers
     source           1..326
                      /organism="Homo sapiens"
                      /note="human"
     STS              127..326
                      /map="760_B_?; (781-788)_C_4; (949-956)_D_11"
     primer_bind      127..151
                      /map="760_B_?; (781-788)_C_4; (949-956)_D_11"
     primer_bind      complement(306..326)
                      /map="760_B_?; (781-788)_C_4; (949-956)_D_11"
BASE COUNT      110 a     42 c     53 g    118 t      3 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 108;  Length 326;
  Best Local Similarity  62.5%;  Pred. No. 3.47e-01;
  Matches    15;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db      9 tattctgtacatattataggaact 32
          |||||          ||||||||||
Qy      9 tattcNNNNNNNNgtataggaact 32


RESULT   31
LOCUS       T40922        345 bp    mRNA            EST       08-FEB-1995
DEFINITION  ya14c01.s1 Homo sapiens cDNA clone 61440 3' similar to
```



```
            SP:C4BB_HUMAN P20851 C4B-BINDING PROTEIN BETA CHAIN ,.
ACCESSION   T40922
KEYWORDS    EST.
SOURCE      human clone=61440 library=Stratagene liver (#937224)
            vector=pBluescript SK host=SOLR cells (kanamycin resistant)
            primer=-21m13 Rsite1=EcoRI Rsite2=XhoI Cloned unidirectionally.
            Primer: Oligo dT. Hepatectomy from normal 49 year old male
            caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; 5'
            adaptor sequence: 5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: ya14c01.r2
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 648505
FEATURES              Location/Qualifiers
     source           1..345
                      /organism="Homo sapiens"
                      /clone="61440"
                      /note="human"
BASE COUNT       84 a     89 c     60 g    112 t
ORIGIN

  Query Match           53.8%;  Score 14;  DB 73;  Length 345;
  Best Local Similarity  58.8%;  Pred. No. 3.47e-01;
  Matches    20;  Conservative    0;  Mismatches  14;  Indels   0;  Gaps   0;

Db    203 gaagtgccttctcacactctggtttgggagcttc 236
          ||||| |||  ||        || | ||| ||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaacttc 34

RESULT   32
LOCUS       ATTS3932      347 bp    RNA             EST       16-SEP-1994
DEFINITION  A. thaliana transcribed sequence; clone FAFM89.
ACCESSION   Z37189
KEYWORDS    expressed sequence tag; partial cDNA sequence.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryotae; mitochondrial eukaryotes; Chlorophyta/Embryophyta
            group; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Magnoliopsida; Dilleniidae; Capparales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 347)
  AUTHORS   Parmentier,Y., Criqui,M.C., Durr,A. and Fleck,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1994) to the EMBL/GenBank/DDBJ databases. CNRS,
            GDR-1003 ACS, INRA, laboratoire de Biologie Moleculaire, BP 27,
```

```
            31326 Castanet-Tolosan cedex, France.
            E-mail:gdr-svp@toulouse.inra.fr. On behalf of: Laboratoire de
            Biologie Moleculaire des Plantes - CNRS, Fleck Jacqueline / L626,
            12 Rue du General Zimmer, 67084 Strasbourg Cedex, France.
            E-mail:fleck@scilla.u-strasbg.fr
REFERENCE   2  (bases 1 to 347)
  AUTHORS   CNRS.
  TITLE     The Arabidopsis thaliana transcribed genome: the GDR cDNA program
  JOURNAL   Unpublished
COMMENT     Cloning vector: Lambda ZAPII;
            Physiological condition: leaves strips incubated 2/3/4 days
                                     in liquid culture medium. full automatic.

            NCBI gi: 586962
FEATURES             Location/Qualifiers
     source          1..347
                     /organism="Arabidopsis thaliana"
                     /clone="FAFM89"
                     /tissue_type="sliced leaves of A.thaliana ecotype
                     columbia"
                     /clone_lib="Strasbourg-A"
BASE COUNT      102 a     54 c     98 g     93 t
ORIGIN

  Query Match           53.8%;  Score 14;  DB 1;  Length 347;
  Best Local Similarity  61.5%;  Pred. No. 3.47e-01;
  Matches    16;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db     89 gaagttgctaaacagattgaagaata 114
          |||||| ||| ||        |||||
Cp     34 gaagttcctatacNNNNNNNNgaata 9


RESULT   33
LOCUS       HSC30G121     350 bp    RNA             EST       04-MAR-1995
DEFINITION  H. sapiens partial cDNA sequence; clone c-30g12.
ACCESSION   F11729
KEYWORDS    partial cDNA sequence; transcribed sequence fragment.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Genexpress.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-1995) to the EMBL/GenBank/DDBJ databases.
            Genethon, B.P. 60, 91002 Evry Cedex France and Genetique
            Moleculaire et Biologie du developpement, CNRS UPR420 B.P. 8, 94801
            Villejuif Cedex France. E-mail: genexpress@genethon.fr
REFERENCE   2  (bases 1 to 350)
  AUTHORS   Genexpress.
  TITLE     The Genexpress cDNA program
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 350)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Boucher,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,
            Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: Integated molecular analysis of the human genome and its
            expression
  JOURNAL   C.R. Acad. Sci., III, Sci. Vie 318, 263-272 (1995)
```



```
COMMENT     Cloning_method: total mRNA was oligo-(dT) primed and directionally
            cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA
            vector;
            Sequencing_method: single read, full automatic;
            Primer: M13_reverse
            cDNA sequence colinear to mRNA
            Stretch_removed: nothing
            Normalization_method: Bento Soares, P.N.A.S. 91:9228-9232(1994);
            Genexpress_library_idt: C;
            Genexpress_sequence_idt: y1c-30g12.

            NCBI gi: 706035
FEATURES             Location/Qualifiers
     source          1..350
                     /organism="Homo sapiens"
                     /clone_lib="normalized infant brain cDNA from B.Soares,
                     Psychiatry Dept. Columbia University USA"
                     /sex="female"
                     /tissue_type="total brain"
                     /dev_stage="3 months old"
                     /isolate="muscular atrophy patient"
BASE COUNT      103 a     88 c     83 g     75 t      1 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 27;  Length 350;
  Best Local Similarity  60.0%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    237 gttcccattctgtgtgcagcctagggactt 266
          ||||| ||||        |  |||| ||||
Qy      4 gttcctattcNNNNNNNNgtataggaactt 33


RESULT   34
ID   HS113E01A  standard; RNA; EST; 369 BP.
AC   D60505;
DT   27-AUG-1995 (Rel. 45, Created)
DT   27-AUG-1995 (Rel. 45, Last updated, Version 1)
DE   Human fetal brain cDNA 3'-end GEN-113E01.
KW   EST(expressed sequence tag); Human fetal brain;
KW   similar to none(May 29,1995).
OS   Homo sapiens (human)
OC   Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC   Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
RN   [1]
RP   1-369
RA   Fujiwara T., Hirano H., Katagiri T., Kawai A., Kuga Y., Nagata M.,
RA   Okuno S., Ozaki K., Shimizu F., Shimada Y., Shinomiya H.,
RA   Takaichi A., Takeda S., Watanabe T., Takahashi E.i., Hirai Y.,
RA   Maekawa H., Shin S., Nakamura Y.;
RT   "Unpublished(101)";
RL   Unpublished.
CC   Submitted (30-May-1995) to DDBJ by: Tsutomu Fujiwara Otsuka GEN
CC   Research Institute Otsuka Pharmaceutical Co.,Ltd 463-10 kagasuno
CC   Kawauchi-cho Tokushima, Tokushima 771-01 Japan Phone: 0886-65-2888
CC   Fax : 0886-37-1035
FH   Key             Location/Qualifiers
FH
FT   source          1..369
FT                   /organism="Homo sapiens"
FT                   /sequenced_mol="cDNA to mRNA"
FT                   /clone_lib="Clontech human fetal brain polyA+ mRNA (#6535)
SQ   Sequence 369 BP; 154 A; 57 C; 59 G; 90 T; 9 other;
```

```
  Query Match           53.8%;  Score 14;  DB 114;  Length 369;
  Best Local Similarity  60.0%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    166 gttcccataaaattgtccaaatagaaactt 195
          ||||| |||          ||||| |||||
Cp     31 gttcctatacNNNNNNNNgaataggaactt 2


RESULT   35
LOCUS       R69283        375 bp    mRNA            EST       01-JUN-1995
DEFINITION  yi39a11.s1 Homo sapiens cDNA clone 141596 3'.
ACCESSION   R69283
KEYWORDS    EST.
SOURCE      human clone=141596 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=Promega -21m13 Rsite1=Not I Rsite2=Eco RI Female
            placenta obtained at birth (full term). 1st strand cDNA was primed
            with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 375)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 273
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 842800
FEATURES              Location/Qualifiers
     source           1..375
                      /organism="Homo sapiens"
                      /clone="141596"
                      /note="human"
BASE COUNT      110 a      55 c      79 g     124 t       7 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 53;  Length 375;
  Best Local Similarity  60.0%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;
```



```
Db    192 gaagttcttattgtacatgtggcatgggaa 221
          ||||||| ||||         | || ||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaa 30


RESULT   36
LOCUS       R27838        381 bp    mRNA            EST       25-APR-1995
DEFINITION  yh65h04.s1 Homo sapiens cDNA clone 134647 3'.
ACCESSION   R27838
KEYWORDS    EST.
SOURCE      human clone=134647 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=-21m13 Rsite1=Not I Rsite2=Eco RI Female placenta
            obtained at birth (full term). 1st strand cDNA was primed with a
            Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 381)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 194
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 783973
FEATURES              Location/Qualifiers
     source           1..381
                      /organism="Homo sapiens"
                      /clone="134647"
                      /note="human"
BASE COUNT      117 a      65 c      66 g     126 t       7 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 41;  Length 381;
  Best Local Similarity  59.4%;  Pred. No. 3.47e-01;
  Matches    19;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db    168 aagttcctctattattcatagaaaagaaattt 199
          |||||||| ||         ||| || || ||
Cp     33 aagttcctatacNNNNNNNNgaataggaactt 2


RESULT   37
LOCUS       R11119        390 bp    mRNA            EST       11-APR-1995
```

```
DEFINITION  yf39g03.r1 Homo sapiens cDNA clone 129268 5' similar to
            SP:TOA2_YEAST P32774 TRANSCRIPTION INITIATION FACTOR IIA SMALL
            CHAIN ;.
ACCESSION   R11119
KEYWORDS    EST.
SOURCE      human clone=129268 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsitel=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 267
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 763854
FEATURES              Location/Qualifiers
     source           1..390
                      /organism="Homo sapiens"
                      /clone="129268"
                      /note="human"
BASE COUNT      114 a     73 c     83 g    115 t      5 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 36;  Length 390;
  Best Local Similarity  62.5%;  Pred. No. 3.47e-01;
  Matches    15;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    275 tccaatactacagaatgaatagga 298
          ||| ||||       ||||||||
Cp     29 tcctatacNNNNNNNNgaatagga 6


RESULT   38
LOCUS       R28016        418 bp    mRNA            EST       25-APR-1995
DEFINITION  yh58f05.s1 Homo sapiens cDNA clone 133953 3'.
ACCESSION   R28016
KEYWORDS    EST.
SOURCE      human clone=133953 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
```



```
            resistant) primer=Promega -21m13 Rsitel=Not I Rsite2=Eco RI Female
            placenta obtained at birth (full term). 1st strand cDNA was primed
            with a Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 418)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 333
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 784151
FEATURES              Location/Qualifiers
     source           1..418
                      /organism="Homo sapiens"
                      /clone="133953"
                      /note="human"
BASE COUNT      132 a     60 c     85 g    137 t      4 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 41;  Length 418;
  Best Local Similarity  60.0%;  Pred. No. 3.47e-01;
  Matches    18;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    167 gaagttcttattgtacatgtagcatgggaa 196
          ||||||| ||||        | || ||||
Qy      1 gaagttcctattcNNNNNNNNgtataggaa 30


RESULT   39
LOCUS       R01221        421 bp    mRNA            EST       31-MAR-1995
DEFINITION  ye81a01.s1 Homo sapiens cDNA clone 124104 3'.
ACCESSION   R01221
KEYWORDS    EST.
SOURCE      human clone=124104 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=-21m13 Rsitel=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
```

```
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 361
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 750957
FEATURES              Location/Qualifiers
     source           1..421
                      /organism="Homo sapiens"
                      /clone="124104"
                      /note="human"
BASE COUNT      129 a      72 c      74 g     142 t       4 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 34;  Length 421;
  Best Local Similarity  60.0%;  Pred. No. 3.47e-01;
  Matches    15;  Conservative    0;  Mismatches  10;  Indels   0;  Gaps   0;

Db    255 ttnctattcattctgtagtattgga 279
          || ||||||        |||| |||
Qy      5 ttcctattcNNNNNNNNgtataggа 29


RESULT   40
LOCUS       T71405        450 bp    mRNA            EST       15-MAR-1995
DEFINITION  yd35c09.r1 Homo sapiens cDNA clone 110224 5' similar to
            gb|L13707|HUMSCALUE Human scRNA molecule, transcribed from Alu
            (rRNA); gb:J04970 CARBOXYPEPTIDASE M PRECURSOR (HUMAN);contains Alu
            repetitive element;.
ACCESSION   T71405
KEYWORDS    EST.
SOURCE      human clone=110224 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
```



```
REFERENCE   1  (bases 1 to 450)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: yd35c09.s1
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 327
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 685926
FEATURES              Location/Qualifiers
     source           1..450
                      /organism="Homo sapiens"
                      /clone="110224"
                      /note="human"
BASE COUNT      111 a      93 c     110 g     133 t       3 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 81;  Length 450;
  Best Local Similarity  62.5%;  Pred. No. 3.47e-01;
  Matches    15;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    145 attctttatttggtattggaactt 168
          ||||        |||| |||||||
Qy     10 attcNNNNNNNNgtataggaactt 33


RESULT   41
LOCUS       H44838        468 bp    mRNA            EST       31-JUL-1995
DEFINITION  yp24g12.s1 Homo sapiens cDNA clone 188422 3'.
ACCESSION   H44838
KEYWORDS    EST.
SOURCE      human clone=188422 library=Soares breast 3NbHBst vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=Promega -21m13 Rsite1=Not I Rsite2=Eco RI Adult
            human. 1st strand cDNA was primed with a Not I - oligo(dT) primer
            [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
            digested with Not I and cloned into the Not I and Eco RI sites of a
            modified pT7T3 vector (Pharmacia). Library went through one round
            of normalization to a Cot = 20. Library constructed by Bento Soares
            and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
```

```
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 920890
FEATURES              Location/Qualifiers
     source           1..468
                      /organism="Homo sapiens"
                      /clone="188422"
                      /note="human"
BASE COUNT      121 a    109 c    101 g    133 t      4 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 17;  Length 468;
  Best Local Similarity  60.7%;  Pred. No. 3.47e-01;
  Matches    17;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db     65 cccattccagcctgcttattggaacttc 92
          || ||||         ||| ||||||||
Qy      7 cctattcNNNNNNNNgtataggaacttc 34


RESULT   42
LOCUS       T51605        471 bp    mRNA            EST       08-FEB-1995
DEFINITION  yb27g03.s1 Homo sapiens cDNA clone 72436 3'.
ACCESSION   T51605
KEYWORDS    EST.
SOURCE      human clone=72436 library=Stratagene fetal spleen (#937205)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=-21m13 Rsite1=EcoRI Rsite2=XhoI Pooled fetal spleens. Cloned
            unidirectionally.  Primer: Oligo dT. Average insert size: 1.0 kb;
            Uni-ZAP XR Vector; 5' adaptor sequence: 5'-GAATTCGGCACGAG-3'; 3'
            adaptor sequence: 5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 380
```



```
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 653465
FEATURES              Location/Qualifiers
     source           1..471
                      /organism="Homo sapiens"
                      /clone="72436"
                      /note="human"
BASE COUNT      143 a     81 c     95 g    150 t      2 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 76;  Length 471;
  Best Local Similarity  59.4%;  Pred. No. 3.47e-01;
  Matches    19;  Conservative    0;  Mismatches  13;  Indels   0;  Gaps   0;

Db    174 aagttcctctattattcatagaaaagaaattt 205
          |||||||| ||         ||| || || ||
Cp     33 aagttcctatacNNNNNNNNgaataggaactt 2


RESULT   43
LOCUS       H42585        473 bp    mRNA            EST       31-JUL-1995
DEFINITION  yo09d08.r1 Homo sapiens cDNA clone 177423 5'.
ACCESSION   H42585
KEYWORDS    EST.
SOURCE      Homo sapiens clone=177423 library=Soares adult brain N2b5HB55Y
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13Rev Rsite1=Not I Rsite2=Eco RI
            55-year old male. 1st strand cDNA was primed with a Not I -
            oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). Library went
            through one round of normalization to a Cot = 53. Library
            constructed by Bento Soares and M.Fatima Bonaldo. The adult brain
            RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18
            hours after death which occurred in consequence of a ruptured
            aortic aneurysm. RNA was prepared from a pool of tissues
            representing the following areas of the brain: frontal, parietal,
            temporal and occipital cortex from the left and right hemispheres,
            subcortical white matter, basal ganglia, thalamus, cerebellum,
            midbrain, pons and medulla.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 169
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 918637
FEATURES              Location/Qualifiers
     source           1..473
                      /organism="Homo sapiens"
                      /clone="177423"
BASE COUNT      119 a      79 c      98 g     162 t      15 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 16;  Length 473;
  Best Local Similarity  60.7%;  Pred. No. 3.47e-01;
  Matches    17;  Conservative    0;  Mismatches  11;  Indels   0;  Gaps   0;

Db    136 gttcttattcatatatttttataggcac 163
          |||| |||||          |||||| ||
Qy      4 gttcctattcNNNNNNNNgtataggaac 31


RESULT   44
LOCUS       H01413        476 bp    mRNA            EST       19-JUN-1995
DEFINITION  yi99c10.r1 Homo sapiens cDNA clone 147378 5'.
ACCESSION   H01413
KEYWORDS    EST.
SOURCE      human clone=147378 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Female placenta
            obtained at birth (full term). 1st strand cDNA was primed with a
            Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 476)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 339
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
```



```
            NCBI gi: 864346
FEATURES              Location/Qualifiers
     source           1..476
                      /organism="Homo sapiens"
                      /clone="147378"
                      /note="human"
BASE COUNT      137 a      77 c      78 g     177 t       7 others
ORIGIN

  Query Match           53.8%;  Score 14;  DB 4;  Length 476;
  Best Local Similarity  65.4%;  Pred. No. 3.47e-01;
  Matches    17;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    386 gttcctatgcntatttttttaatagga 411
          |||||||| ||        |||||||
Cp     31 gttcctatacNNNNNNNNgaatagga 6


RESULT   45
LOCUS       T59688        516 bp    mRNA            EST       09-FEB-1995
DEFINITION  yc13f05.s1 Homo sapiens cDNA clone 80577 3' similar to
            gb:D90150_rna1 GUANINE NUCLEOTIDE-BINDING PROTEIN G(Z), ALPHA
            SUBUNIT (HUMAN);.
ACCESSION   T59688
KEYWORDS    EST.
SOURCE      human clone=80577 library=Stratagene lung (#937210)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=-21m13 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72
            year old male. Cloned unidirectionally. Primer: Oligo dT. Average
            insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
            5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 324
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

            NCBI gi: 661525
FEATURES              Location/Qualifiers
     source           1..516
                      /organism="Homo sapiens"
                      /clone="80577"
                      /note="human"
BASE COUNT      148 a     110 c     102 g     151 t       5 others
ORIGIN
```

```
  Query Match           53.8%;  Score 14;  DB 78;  Length 516;
  Best Local Similarity  57.6%;  Pred. No. 3.47e-01;
  Matches    19;  Conservative    0;  Mismatches  14;  Indels   0;  Gaps   0;

Db    311 gaaatttatatagggtgaggacaataggaantt 343
          ||| ||  ||||          |||||||| ||
Cp     34 gaagttcctatacNNNNNNNNgaataggaactt 2


Search completed: Tue May 14 14:08:43 1996
Job time : 526 secs.
```

Set	Items	Description
S1	1	AU=WAHL ? AND FLP
S2	0	O'GORMAN S? AND FLP
S3	1	AU=O'GORMAN S? AND FLP
S4	0	AU="O'GORMAN S?" AND FLP
S5	1	FLP AND MICE
S6	9	(FRT OR FLP) AND TRANSGEN?

5,510,099

?t6/3/1-9

6/3/1
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11994591 BIOSIS Number: 98594591
Genetic regulation of mec-3 gene expression implicated in the specification of the mechanosensory neuron cell types in Caenorhabditis elegans
Mitani S
Dep. Physiol., Tokyo Women's Med. Coll., 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162, Japan
Development Growth & Differentiation 37 (5). 1995. 551-557.
Full Journal Title: Development Growth & Differentiation
ISSN: 0012-1592
Language: ENGLISH
Print Number: Biological Abstracts Vol. 101 Iss. 002 Ref. 022296

6/3/2
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11963067 BIOSIS Number: 98563067
FLP recombinase in transgenic plants: Constitutive activity in stably transformed tobacco and generation of marked cell clones in Arabidopsis
Kilby N J; Davies G J; Snaith M R; Murray J A H
Inst. Biotechnol., Univ. Cambridge, Tennis Court Rd., Cambridge CB2 1QT, UK
Plant Journal 8 (5). 1995. 637-652.
Full Journal Title: Plant Journal
ISSN: 0960-7412
Language: ENGLISH
Print Number: Biological Abstracts Vol. 101 Iss. 001 Ref. 005872

6/3/3

DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11947956 BIOSIS Number: 98547956
Activity of the yeast FLP recombinase in Arabidopsis
Sonti R V; Tissier A F; Wong D; Viret J-F; Signer E R
Dep. Biol., Mass. Inst. Technol., Cambridge, MA 02138-4307, USA
Plant Molecular Biology 28 (6). 1995. 1127-1132.
Full Journal Title: Plant Molecular Biology
ISSN: 0167-4412
Language: ENGLISH
Print Number: Biological Abstracts Vol. 100 Iss. 012 Ref. 184647

6/3/4
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11852250 BIOSIS Number: 98452250
Heat-inducible expression of FLP gene in maize cells
Lyznik L A; Hirayama L; Rao K V; Abad A; Hodges T K
Dep. Botany and Plant Pathol., Purdue Univ., West Lafayette, IN 47907, USA
Plant Journal 8 (2). 1995. 177-186.
Full Journal Title: Plant Journal
ISSN: 0960-7412
Language: ENGLISH
Print Number: Biological Abstracts Vol. 100 Iss. 008 Ref. 118307

6/3/5
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11521508 BIOSIS Number: 98121508
Site-specific Transgene Insertion: An Approach
Wigley P; Becker C; Beltrame J; Blake T; Crocker L; Harrison S; Lyons I; McKenzie Z; Tearle R; Crawford R; Robins A
Bresatec Lab., Dep. Biochem., Univ. Adelaide, SA 5005, Australia
Reproduction Fertility and Development 6 (5). 1994. 585-588.
Full Journal Title: Reproduction Fertility and Development
ISSN: 1031-3613
Language: ENGLISH
Print Number: Biological Abstracts Vol. 099 Iss. 006 Ref. 078065

6/3/6

DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11204786 BIOSIS Number: 97404786
A cell-autonomous, ubiquitous marker for the analysis of Drosophila genetic mosaics
Vincent J-P; Girdham C H; O'Farrell P H
Dep. Biochem. Biophysics, Univ. California at San Francisco, San Francisco, CA 94143-0048, USA
Developmental Biology 164 (1). 1994. 328-331.
Full Journal Title: Developmental Biology
ISSN: 0012-1606
Language: ENGLISH
Print Number: Biological Abstracts Vol. 098 Iss. 006 Ref. 075598

6/3/7
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

10922464 BIOSIS Number: 97122464
A new method for manipulating transgenes: Engineering heat tolerance in a complex, multicellular organism
Welte M A; Tetrault J M; Dellavalle R P; Lindquist S L
Howard Hughes Med. Inst., Univ. Chicago, 5841 South Maryland, MC 1028 Chicago, IL 60637, USA
Current Biology 3 (12). 1993. 842-853.
Full Journal Title: Current Biology
ISSN: 0960-9822
Language: ENGLISH
Print Number: Biological Abstracts Vol. 097 Iss. 006 Ref. 072338

6/3/8
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

9042030 BIOSIS Number: 93027030
FLP-MEDIATED RECOMBINATION IN THE VECTOR MOSQUITO AEDES-AEGYPTI
MORRIS A C; SCHAUB T L; JAMES A A
DEP. MOL. BIOL. AND BIOCHEM., UNIV. CALIF., IRVINE, CALIF. 92717.
NUCLEIC ACIDS RES 19 (21). 1991. 5895-5900. CODEN: NARHA
Full Journal Title: Nucleic Acids Research
Language: ENGLISH

6/3/9

6/7/5
DIALOG(R)File 5:BIOSIS PREVIEWS(R)

11521508 BIOSIS Number: 98121508
Site-specific Transgene Insertion: An Approach
Wigley P; Becker C; Beltrame J; Blake T; Crocker L; Harrison S; Lyons I; McKenzie Z; Tearle R; Crawford R; Robins A
Bresatec Lab., Dep. Biochem., Univ. Adelaide, SA 5005, Australia
Reproduction Fertility and Development 6 (5). 1994. 585-588.
Full Journal Title: Reproduction Fertility and Development
ISSN: 1031-3613
Language: ENGLISH
Print Number: Biological Abstracts Vol. 099 Iss. 006 Ref. 078065
Methods to improve the production of transgenic animals are being developed. Conventional transgenesis, involving microinjection of DNA into fertilized eggs, has a number of limitations. These result from the inability to control both the site of transgene insertion and the number of gene copies inserted. The approach described seeks to overcome these problems and to allow single copy insertion of transgenes into a defined site in animal genomes. The method involves the use of embryonic stem cells, gene targeting and the FLP recombinase system.

DIALOG(R)File 5:BIOSIS PREVIEWS(R)

8208682 BIOSIS Number: 91129682
RECOMBINASE-MEDIATED GENE ACTIVATION AND SITE-SPECIFIC INTEGRATION IN MAMMALIAN CELLS
O'GORMAN S; FOX D T; WAHL G M
GENE EXPRESSION LAB., SALK INST. BIOL. STUDIES, LA JOLLA, CALIF. 92037.
SCIENCE (WASHINGTON D C) 251 (4999). 1991. 1351-1355. CODEN: SCIEA
Full Journal Title: SCIENCE (Washington D C)
Language: ENGLISH
?

227: 680-685.

13. Lee, J., and M. Jayaram; 1993. Mechanism of site-specific recombination: logic of assembling recombinase catalytic site from fractional active sites. J. Biol. Chem. 268:17564-17579.

14. Maniatis, T., E. F. Fritsch, and J. Sambrook; 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

15. Nunes-Duby, S. E., L. Matsumoto, and A. Landy; 1987. Site-specific recombination intermediates trapped with suicide substrates. Cell 50:779-788.

16. Nunes-Duby, S. E., R. S. Tirumalai, L. Dorgai, E. Yagil, R. A. Weisberg, and A. Landy; (lambda) integrase cleaves DNA in cis. EMBO J., in press.

17. Pan, G., K. Luetke, and P. D. Sadowski; 1993. Mechanism of cleavage and ligation by the Flp recombinase: classification of mutations in the Flp protein using in vitro complementation analysis. Mol. Cell. Biol. 13:3167-3175.

18. Parsons, R. L., B. R. Evans, L. Zheng, and M. Jayaram; 1990. Functional analysis of Arg-308 mutants of Flp recombinase. J. Biol. Chem. 265:4527-4533.

19. Parsons, R. L., P. V. Prasad, R. M. Harshey, and M. Jayaram; 1988. Step-arrest mutants of FLP recombinase: implications for the catalytic mechanism of DNA recombination. Mol. Cell. Biol. 8:3303-3310.

20. Prasad, P. V., L. J. Young, and M. Jayaram; 1987. Mutations in the 2 (mu)m circle site-specific recombinase that abolish recombination without affecting substrate recognition. Proc. Natl. Acad. Sci. USA 84:2189-2193.

21. Qian, X. H., R. B. Inman, and M. M. Cox; 1990. Protein-based asymmetry and protein-protein interactions in Flp recombinase-mediated site-specific recombination. J. Biol. Chem. 265:4527-4533.

22. Serre, M. C., B. R. Evans, H. Araki, Y. Oshima, and M. Jayaram; 1992. Half-site recombinations mediated by yeast site-specific recombinases Flp and R. J. Mol. Biol. 225:621-642.

23. Serre, M. C., and M. Jayaram; 1992. Half-site transfer by step-arrest mutants of yeast site-specific recombinase Flp. J. Mol. Biol. 225:643-649.

24. Yang, S. H., and M. Jayaram; 1994. Generality of the shared active site among yeast family site-specific recombinases: the R site-specific recombinase follows the Flp paradigm. J. Biol. Chem. 269: 12789-12796.

1. Abremski, K., and R. H. Hoess; 1992. Evidence for a second conserved arginine in the integrase family of recombination proteins. Protein Eng. 5:87-91.

2. Amin, A., H. Roca, K. Luetke, and P. D. Sadowski; 1991. Synapsis, strand scission, and strand exchange induced by the FLP recombinase: analysis with half-FRT sites. Mol. Cell. Biol. 11:4497-4508.

3. Argos, P., A. Landy, K. Abremski, J. B. Egan, E. Haggard-Ljungquist, R. H. Hoess, M. L. Kahn, B. Kalionis, S. V. L. Narayana, L. S. Pierson, III, N. Sternberg, and J. M. Leong; 1986. The integrase family of site-specific recombinases: regional similarity and global diversity. EMBO J. 5:433-440.

4. Beaucage, S. L., and M. H. Caruthers; 1981. Deoxynucleoside phosphoramidites: a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. 22:1859-1862.

5. Blakely, G., G. May, R. McCulloch, L. K. Arciszewska, M. Burke, S. T. Lovett, and D. J. Sherratt; 1993. Two related recombinases are required for site-specific recombination at dif and cer in E. coli K12. Cell 75:351-361.

6. Chen, J. W., J. Lee, and M. Jayaram; 1992. DNA cleavage in trans by the active site tyrosine during Flp recombination: switching protein partners before exchanging strands. Cell 69:647-658.

7. Chen, J. W., S. H. Yang, and M. Jayaram; 1993. Tests for the fractional active-site model in Flp site-specific recombination: assembly of a functional recombination complex in half-site and full-site strand transfer. J. Biol. Chem. 268:14417-14425.

8. Doolittle, R. F.; 1994. Convergent evolution: the need to be explicit. Trends Biochem. Sci. 19:15-18.

9. Evans, B. R., J. W. Chen, R. L. Parsons, T. K. Bauer, D. B. Teplow, and M. Jayaram; 1990. Identification of the active site tyrosine of Flp recombinase: the possible relevance of its location to the mechanism of recombinase. J. Biol. Chem. 265:18504-18510.

10. Han, Y. W., R. I. Gumport, and J. F. Gardner; 1993. Complementation of bacteriophage lambda integrase mutants: evidence for an intersubunit active site. EMBO J. 12:4577-4584.

11. Jayaram, M.; 1994. Mechanism of site-specific recombination: the Flp paradigm. Nucleic Acids Mol. Biol. 8:268-286.

12. Laemmli, U. K.; 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London)

3(prm1)TC5(prm1)) adjacent to the top-strand cleavage point. The unlabeled substrate was a concatemer of S2 obtained by self-ligation (S2n). S2 also contained two mismatched spacer positions next to the top-strand cleavage site (5(prm1)AG3(prm1)/3(prm1)AA5(prm1)). Following strand swapping between S1 and S2, the spacers would be fully matched (5(prm1)AG3(prm1)/3(prm1)TC5(prm1) and 5(prm1)TT3(prm1)/3(prm1)AA5(prm1)). Strand cleavage and strand transfer products are designated CP and STP, respectively. S represents the substrate. The heterogeneity in strand transfer products results from the multiplicity of crossover points within S2n. For each reaction set with a complementing protein pair, the leftmost lane represents a reaction with the triad variant alone (at the same molar concentration as in the rightmost lane). The next three lanes represent reactions containing the triad variant and Flp(Y343F) in approximate molar ratios of 1:1, 1:1.5, and 1:2, respectively. Roughly 3 pmol of Flp(Y343F) was present per pmol of the Flp-binding element. Lane C represents an assay with no Flp or Flp variant added to the reaction. The Flp reaction shown in lane 2 contained S1 but not S2n. The product X likely arose by cleavage within S1 and subsequent phosphoryl transfer to the 5(prm1)-hydroxyl group of the unannealed top-strand oligodeoxynucleotide of S1 present in the reaction. The size of X, as measured against standard molecular size markers, fits this explanation. Note the presence of X in a reaction containing S1 alone (lane 2). WT, wild type.

End_Page 7497----------

recombination by Flp, they do not allow us to distinguish among the three potential types of trans DNA cleavage, trans horizontal, trans vertical, and trans diagonal (6; Fig. 1A). Results obtained with half-site reactions tend to disfavor the trans-vertical mode, while distinction between the trans-horizontal and trans-diagonal modes is not possible. Our expressed preference for trans-diagonal cleavage (6) over trans-horizontal cleavage must be tempered by the possibility that half sites are likely to enjoy greater freedom of stacking interactions in solution over full sites (11; Fig. 1B). The critical question is: what is the cleavage mode in full-site recombination? The answer can be sought provided a tagged recombinase can be targeted to a specific binding arm within a full-site recombination complex.

ACKNOWLEDGMENTS

We thank S. E. Nunes-Duby and A. Landy for communicating results prior to publication.

This work was supported by a grant from the N.I.H. Partial support was received from the Council for Tobacco Research and the Robert F. Welch Foundation.

References:

Flp assembles a functional active site from partial active sites during full-site recombination:
The apparent cis DNA cleavage by Int in full attB sites and Holliday junctions calls into question the pertinence of the active-site assembly by Flp during a half-site reaction to that during normal recombination. The design of full sites containing mismatched spacers has allowed us to address this issue directly. Complementation by step arrest variants of Flp to mediate strand cleavage in full sites supports the shared-active-site configuration during a normal recombination reaction. In a reaction carried out by a pair of Flp step arrest mutants, cleavage is executed by the protein partner harboring Tyr-343, thus virtually excluding the operation of an aberrant pathway during complementation.

Recently, mechanistic analyses of another Int family recombinase, Xer (responsible for stable chromosome partitioning in Escherichia coli), have become possible (5). Recombination in this system requires the combined action of two recombinases, XerC and XerD. The binding arms of the Xer target site encode specificities for each of the two protein monomers. The cleavage pattern observed with Xer is most easily explained by cis DNA cleavage, although one particular type of trans cleavage (trans vertical or partner trans; 6, 16) cannot be ruled out.

The (lambda) Int and Xer examples, contrasted with Flp, imply that Int family recombinases do not conform to the same rules in building their active sites. However, one suspects that, within the fully assembled active sites of these proteins, key catalytic residues must have the same relative spatial disposition. This would account for the fact that they follow the same chemistry of recombination. Global diversity and limited homology, which are the hallmarks of this family (3), would then make a strong case for mechanistic convergence (8, 16) among proteins that execute chemically identical reactions.

Which mode of trans cleavage does Flp follow?:
While our results strongly support trans DNA cleavage during full-site

FIG. 5. [GREY SCALE PLATE AVAILABLE] Complementation between Flp(Y343F) and double or triple triad mutants of Flp in full-site cleavage. The substrate was labeled at the 3(prm1) ends. The proteins used in the assays are indicated above the lanes. The substrate band (S) and the cleavage products (CL and CR) are labeled as in Fig. 3. A control reaction without Flp or Flp variants is shown in lane C.

FIG. 6. [GREY SCALE PLATE AVAILABLE] Strand transfer in full sites by pairwise combinations of Flp step arrest mutants. The substrates used in the assay are schematically shown at the top. The labeled substrate S1 contained (sup32)P at the 3(prm1) end of the top strand and two spacer mismatches (5(prm1)TT3(prm1) and

The successful execution of strand exchange within a full site by a complementing pair of Flp variants fully corroborates the inference from half-site reactions that Flp(Y343F) (harboring an intact RHR triad) can facilitate not only strand cleavage but also strand joining by using a Tyr-343 residue and a 5(prm1)-hydroxyl group, respectively, donated in trans.

DISCUSSION

The partial-active-site-trans DNA cleavage model for Flp was first proposed to account for the pattern of DNA cleavage and strand transfer in half-site substrates by complementing

FIG. 3. [GREY SCALE PLATE AVAILABLE] Complementation between Flp(Y343F) and triad mutants of Flp in full-site substrates. The full site used in the assays is schematically represented at the top (Fig. 2). The parallel arrows represent the Flp-binding elements; the short vertical arrows indicate the points of strand cleavage by Flp. The two mismatched positions within the spacer adjacent to the cleavage sites are shown by the bubbles. DNA sequences unrelated to recombination are symbolized by the wavy lines. The asterisks stand for the (sup32)P label at the 3(prm1) ends. Products of cleavage from the top strand (left) and the bottom strand (right) are labeled CL and CR, respectively. The substrate band is designated S. The lane marked C is a reaction in which no Flp or Flp variant was added. WT, wild type.

FIG. 4. [GREY SCALE PLATE AVAILABLE] Identification of the protein partner responsible for strand cleavage during catalytic complementation. Strand cleavage assays were carried out by using the substrate used in reactions shown in Fig. 3. The radioactive label was placed at the 5(prm1) end of each strand (asterisks). The covalent DNA-protein complex resulting from strand cleavage by Flp (or a Flp variant) is called DPC; that derived from GST-Flp (or a GST-Flp variant) is called DPCG. The doublets corresponding to DPC or DPCG are most likely cleavage products derived from the top and bottom strands. The substrate DNA band is marked S. Flp and the Flp variants used in the reactions are indicated above the appropriate lanes. Lane C is a control reaction without added Flp or Flp variants. WT, wild type.

End_Page 7496----------

pairs of catalytic Flp mutants (6, 7, 13). The validity of the model was then verified for the R recombinase from Z. rouxii in half-site strand transfer reactions (24). The simplicity and functional parsimony of the model led us to speculate that the rules of active-site assembly and of DNA cleavage are likely to extend beyond the yeast site-specific recombinases and encompass the entire Int family. Experiments with (lambda) Int and suicide attL, attB, or Holliday junction substrates have yielded rather paradoxical results (10, 16). While one set of results supports the Flp paradigm (10), the other set casts doubt on the generality of the model (16).

recombination could be kept hidden. However, difficulty in selectively binding a protein monomer to one of the two normal binding elements of the full-site substrate poses an impediment to the rigorous testing of this prediction. Nevertheless, the degree of cleavage reduction obtained upon the mixing of roughly equal amounts of Flp and Flp(H305L, Y343F) was consistent with inactivation of the wild-type partner in a double-mutant-wild-type protein pair (data not shown).

Strand transfer in bubbled full sites by pairwise combination of Flp(Y343F) and triad variants of Flp:
According to the partial-active-site-trans-cleavage model, during complementation between a Flp triad mutant and Flp(Y343F), cleavage should occur on the scissile phosphodiester adjacent to the bound Flp(Y343F) (the cis configuration in Fig. 1A) and away from the triad mutant (the trans configuration in Fig. 1A) (6, 7). Once the DNA 3(prm1)-phosphotyrosyl bond has been formed, Flp(Y343F) can facilitate the strand-joining reaction by using the 5(prm1)-hydroxyl group as the nucleophile (13, 17). In the bubbled full-site substrate, the spacer mismatch inhibits the joining reaction. However, if one used a pair of substrates (S1 and S2n in Fig. 6), each mismatched within its spacer but fully matched with the partner's spacer, strand joining within a substrate (parental mode) would be suppressed but that between partners (recombinant mode) would be encouraged. Even when the lack of a second pair of exchanges (as with a pair of complementing mutants) would tend to reverse this reaction, one might expect to trap some of the strand transfer product. With wild-type Flp, strand transfer products were formed from this pair of substrates (lane 3 in Fig. 6). The heterogeneity of strand transfer products results from the fact that one of the two DNA substrates (the nonradioactive one) was a concatemer of a single full site (S2) obtained by ligation. We resorted to this trick because, under our assay conditions, strand transfer efficiency was increased severalfold by increasing the length of at least one of the recombination partners. The size of the recombinant strand would depend on the site within S2n at which crossover occurred. In reactions containing Flp(Y343F) in combination with a triad single mutant or a triad double mutant, strand transfer was indeed detected (lanes 6 to 8, 10 to 12, 14 to 16, and 18 to 20 in Fig. 6). The low level of reaction compared with the wild type is not surprising. Since a single-strand transfer requires a matched pair of cleavage events within the two substrates (on the top strands or the bottom strands), only a reaction complex containing two appropriately positioned Flp(Y343F) and triad mutant monomers would be successful in completing an exchange event. Further, unlike, the wild-type reaction, the mutant pair reaction cannot execute the second pair of exchanges that yields the mature recombinant product. One would expect, therefore, that reversal of the first exchange (due to the absence of a second exchange) would be more pronounced in a reaction containing the complementing partners than in the wild-type reaction.

cleavage product corresponded in size to that obtained with GST-Flp (lane 9 in Fig. 4).

The pattern of DNA-protein adducts observed in the complementation reactions demonstrates that cleavage is carried out exclusively by the protein partner that harbors the active-site tyrosine. The simplest interpretation of the results is that it is indeed Tyr-343 that performs strand cleavage. The more complex scenario in which the active species is derived from the triad variant but is a surrogate nucleophile rather than Tyr-343 is not excluded. However, this possibility is strongly discounted by the fact that no complementation was obtained between Flp(Y343F) and a double variant altered at a triad position and Tyr-343 (data not shown).

The donor of Tyr-343 during catalytic complementation can be mutated at all triad positions:
The shared-active-site model for Flp predicts that a Flp variant doubly or triply mutated in the RHR triad would be as competent as the single mutant in catalytic complementation with Flp(Y343F) provided their binding affinities for the target DNA do not differ significantly. This prediction has been verified in half-site recombination (7). We tested two triad double mutants and a triad triple mutant in combination with Flp(Y343F) in the cleavage assay with the bubbled full site (Fig. 5). Individually, neither the triad mutants nor Flp(Y343F) could effect strand cleavage (lanes 3, 4, 6, and 8 in Fig. 5). In contrast, each pair formed by mixing Flp(Y343F) and a triad mutant in roughly equimolar amounts exhibited approximately the same levels of complementation (lanes 5, 7, and 9 in Fig. 5). A second strong prediction of the partial-active-site model is that a triad mutant that also lacks Tyr-343 in combination with wild-type Flp will produce a catalytically inactive protein pair. This prediction could be directly tested in half-site reactions, since one could load the double mutant on a radioactively labeled half site and the wild-type protein on an unlabeled half site and then monitor

FIG. 2. Full-site substrates containing spacer mismatches. The sequences of the synthetic full sites used in the strand cleavage and strand transfer assays are shown. The Flp-binding elements are in boldface. Sequences flanking the Flp target site are represented by wavy lines. The positions of spacer mismatches (bubbles) are indicated. The experiments in which they were used are indicated by the corresponding figure numbers. S2n refers to a concatemer of S2 (8 to 10 monomeric units, on average). The spacer mismatches in S1 and S2 are such that strand swapping between the two substrates (following strand cleavage) would produce perfect complementarity.

End_Page 7495----------

strand transfer only from the labeled substrate upon mixing of the prebound complexes (7). Thus, the background of wild-type

site is cleaved, the short spacer segment on the cleaved strand does not remain stably hydrogen bonded to its complementary sequence. Hence, it is effectively lost from the reaction center by diffusion. The 5(prm1)-hydroxyl group of the spacer on the noncleaved strand can then act as a phosphoryl acceptor to complete a half-site recombination event. Details of half-site reactions have been previously described (for example, see reference 22). Whereas two Flp monomers bound to a full site are restricted in their interactions by spacer connectivity, two half sites, each associated with a Flp monomer, are not subject to this constraint. They could potentially assume the configurations indicated. These correspond to variations of the trans interactions depicted in panel A.

End_Page 7494----------

These results are consistent with the partial active-site model arrived at from half-site reactions (6). According to this model, a mutant pair can build an active site in which the RHR triad is contributed by Flp(Y343F) and Tyr-343 is contributed by the triad mutant. Therefore, strand cleavage becomes possible.

The nucleophile in the cleavage reaction by a pair of complementing Flp variants is Tyr-343:
The partial-active-site model is based on the tacit assumption that strand cleavage is executed by the lone Tyr-343 present within a pair of the complementing protein monomers. The model would break down if, in the absence of Tyr-343, a substitute nucleophile in the form of a serine, threonine, cysteine, or another tyrosine could take up its function. We have ruled out the possibility of cis cleavage by a nucleophile other than Tyr-343 in complementation reactions with half sites (24). In the full-site context, activation of a surrogate nucleophile within Flp(Y343F) as a result of allosteric interactions among the protein monomers is not impossible. To test the surrogate nucleophile hypothesis, the assay was done with a bubbled full site labeled at the 5(prm1) end on both strands and a complementing pair of Flp variants made up of a normal-sized protein and a 30-kDa larger protein partner obtained as a fusion with GST. The protein-DNA adduct formed by Flp and the GST-Flp hybrid upon DNA cleavage can be distinguished by the difference in electrophoretic migration between them (lanes 2 and 3 in Fig. 4). Flp (Y343F) or GST-Flp(Y343F) did not yield the cleavage product, as expected (lanes 4 and 6 in Fig. 4). Flp(H305L) and GST-Flp(H305L) yielded low levels of the cleavage products with the expected mobilities (lanes 5 and 7 in Fig. 4). The low level of strand cleavage by these proteins was as predicted by the results shown in Fig. 3. However, in partnership with Flp (Y343F), they produced elevated levels of cleavage commensurate with catalytic complementation (lanes 8 and 9 in Fig. 4). When the complementing partners were GST-Flp(Y343F) and Flp(H305L), the size of the cleavage product matched that obtained from reactions with Flp (lane 8 in Fig. 4). When the reaction contained the reciprocal pair, Flp(Y343F) and GST-Flp(H305L), the

substrate (6, 17). However, attempts to obtain catalytic complementation between a triad mutant and Flp(Y343F) in full sites have not been successful. This is not surprising. In a recombination complex containing two monomers of each mutant oriented appropriately, a Holliday junction may be formed. Since the junction cannot be resolved into recombinants, the exchange reaction is reversed to restore the parental configuration. Demonstration of complementation therefore required the use of a suicide substrate in which the reaction intermediates could be readily trapped. We discovered that mismatches at certain positions within the strand exchange (spacer) region of a full site can strongly inhibit the strand-joining step of recombination, thus effectively transforming the sites harboring such mismatches into suicide substrates (unpublished data). For example, the substrate shown in Fig. 3 contains two adjacent mismatches each (bubbles) neighboring the cleavage points at the left and right ends of the spacer (Fig. 2). This substrate was cleaved efficiently by wild-type Flp. However, since strand joining was markedly slowed down (unpublished results), the cleavage product accumulated (lane 2 in Fig. 3). As expected, no cleavage was obtained with Flp(Y343F) (lane 3 in Fig. 3). Flp variants in which either of the two arginine residues from the RHR triad (Arg-191 and Arg-308) were changed failed to produce the cleavage product (lanes 4 and 8 in Fig. 3). It is known that Flp variants of the triad histidine can yield cleavage in a full site but are severely diminished in the ability to reseal strands (19). However, in a full site with the double bubble, cleavage by the histidine variants was significantly lowered relative to that obtained with wild-type Flp (compare lanes 6 and 2 in Fig. 3). It is possible that the absence of His-305, combined with the mismatched spacer configuration, perturbs the normal protein interactions that lead to catalysis. The histidine variants are also known to test as cleavage incompetent when provided with half-site substrates (23). When a triad arginine variant of Flp was mixed with Flp(Y343F), they complemented each other, as evidenced by the cleavage detected within the bubbled full site (lanes 5 and 9 in Fig. 3). Complementation was obtained when Flp (Y343F) was paired with the His-305 variant as well. Whereas cleavage with the His-305 variant alone was weak (lane 6 in Fig. 3), the complementing pair yielded much higher levels of cleavage (lane 7 in Fig. 3).

FIG. 1. Full-site and half-site substrates for Flp site-specific recombination. (A) Each full site contains two invertedly oriented Flp-binding elements (parallel arrows) bordering the strand exchange region (spacer). There is a one-to-one association between a binding element and a Flp monomer. Conceptually, a full site can be split into a left half site (L) and a right half site (R). The phosphodiesters involved in recombination between two full sites are indicated (p). The placement of a Flp monomer with respect to these phosphodiesters can be described as cis (a), trans horizontal (b), trans vertical (c), or trans diagonal (d). (B) A half site contains one Flp-binding element and one scissile phosphodiester. When the

mixture containing 250 mM Tris-HCl (pH 7.8), 4% sodium dodecyl sulfate, 40% glycerol, and 300 mM (beta)-mercaptoethanol. Suitable aliquots were heated at 95 deg C for 4 min and fractionated by electrophoresis in 8% polyacrylamide gels (12). The gels were rinsed in distilled water with gentle shaking, dried, and subjected to autoradiography.

General methods:
Restriction enzyme digestions, isolation of plasmid DNA, and other miscellaneous procedures were done as described by Maniatis et al. (14).

RESULTS

The normal Flp reaction uses two double-stranded DNA substrates, each containing a copy of the Flp recombination target sequence (Fig. 1A). These are referred to as full-site substrates. A recombination event between two full sites requires the cooperative action of four Flp monomers and involves the breakage and reformation of four phosphodiester bonds within DNA (two breakage-union steps within each substrate partner). The disposition of a target-bound Flp monomer with respect to the four scissile phosphodiester bonds can be described as cis, trans horizontal, trans vertical, or trans diagonal. Conceptually, a full site can be split into two half sites, a left half site and a right half site. Half-site substrates (Fig. 1B), originally designed for the (lambda) Int reaction (15) and subsequently adapted for the Flp reaction (2, 21, 22), have simplified the mechanistic analysis of site-specific recombination. A half site contains one Flp-binding element, one scissile phosphodiester, and one 5(prm1)-hydroxyl group that can act as a phosphoryl acceptor. Hence, it is capable of undergoing one strand cleavage and one strand-joining reaction, precisely half of the chemistry that a full site undergoes during a normal recombination event. However, while interactions between two Flp monomers within a full site are constrained by the continuous DNA segment between them (the strand exchange region or the spacer), half sites are not subject to such constraints. Hence, two half sites could potentially interact with each other in a variety of modes that may not be accessible to two full sites (Fig. 1B). A legitimate concern, then, is that direct extrapolations from half-site to full-site reactions may not always be valid. To overcome this impediment, the analyses described here were done with appropriately modified full-site substrates (Fig. 2).

Pairwise complementation between Flp(Y343F) and the RHR triad variants of Flp during strand cleavage in full sites:
It is known that a Flp variant altered at one or more of the RHR triad positions in combination with a second variant

lacking the active-site nucleophile, Flp(Y343F), can mediate a strand cleavage and a strand joining event within a half-site

heated to 65 deg C for 10 min, and cooled slowly to room temperature. The relevant features of these sites are described in Results and displayed in the figures. The complete sequences of the substrates are available upon request.

The 3(prm1) end of a deoxyoligonucleotide was labeled with (alpha)-(sup32)P-labeled cordycepin phosphate. Labeling at the 5(prm1) end was done by the T4 polynucleotide kinase reaction by using [(gamma)-(sup32)P]ATP as the phosphoryl donor. For some experiments, the 5(prm1) ends were phosphorylated with unlabeled ATP. The unreacted cordycepin phosphate or ATP was removed by spin dialysis on a G-25 column. Hybridization to the partner oligodeoxynucleotide was done in TE.

Strand cleavage assays:
Strand cleavage reactions were done under standard recombination conditions (6). Normally, 0.05 pmol of the 3(prm1) end-labeled substrate was reacted with approximately 0.5 pmol of Flp or a Flp variant (roughly 5 pmol of Flp per pmol of the binding element) in 30 (mu)l of the reaction mixture. Incubations were done at 30 deg C for 30 min. Reactions were stopped by immersing samples in a boiling water bath for 5 min. After addition of sodium dodecyl sulfate (final concentration, 0.1%) and proteinase K treatment (100 (mu)g per sample for 1 h at 37 deg C), samples were phenol-chloroform extracted and DNA was precipitated with ethanol. The DNA pellet was recovered by centrifugation, washed twice with 80% ethanol, and dried in vacuo. Strands were denatured in 95% formamide at 95 deg C, and samples were fractionated by electrophoresis in 10% denaturing polyacrylamide gels (acrylamide-bisacrylamide ratio, 19:1). Cleavage products were identified following autoradiography.

Strand transfer assays:
The synthetic full sites (approximately 45 to 50 bp long, carrying EcoRI and HindIII overhangs at the ends) were poor substrates in strand transfer. To increase the efficiency of the reaction, assays were done with the monomeric form of a radioactively labeled full site and the concatemeric form of the unlabeled full-site partner. The concatemer was prepared as follows. The full site phosphorylated at the 5(prm1) end on both strands was ligated at room temperature for 3 h under conditions that gave concatemers containing an average of 8 to 10 U of the monomer. The strand transfer reactions were done with the normal protocols described previously (6). The ratio of the labeled substrate to the monomeric equivalent of the unlabeled substrate was approximately 1:5. In these assays, approximately 6 to 8 pmol of Flp per pmol of the Flp-binding element was present in a reaction volume of 30 (mu)l. The samples were processed and fractionated as described for the strand cleavage assay.

Assay for formation of DNA-protein covalent adducts:
Reactions were carried out under strand transfer conditions with a substrate labeled at the 5(prm1) ends on both strands. Reactions were quenched by addition of an equal volume of a stop

The catalytic strategies displayed by Flp and their mechanistic implications in recombination suggest that they may be universal to the Int family. One set of experiments with (lambda) Int using suicide attL substrates supports this notion. Catalytic complementation in pairwise combinations of the RHR triad mutants of Int with the active-site tyrosine mutant has been demonstrated (10). This result is strongly suggestive of trans DNA cleavage by Int. However, other results obtained by using suicide attB substrates and synthetic Holliday junctions are more parsimoniously explained in terms of cis DNA cleavage by Int (16).

The mechanistic dilemma posed by the Int results raises fundamental issues regarding the mechanism of Int family

End_Page 7492----------

site-specific recombination. First, is the apparent cis-trans duality a basic feature of the reaction? Second, are there multiple modes of active-site assembly within this family? Finally, is the half-site reaction mechanistically distinct from the full-site reaction? We have devised an experimental design in which full sites carrying mismatches in the spacer region serve as substrates in complementation tests with step arrest Flp mutants. Our results fully support the shared active-site paradigm during full-site recombination. Furthermore, the mode of DNA cleavage is trans. We have found no evidence of cis-trans duality in Flp recombination.

MATERIALS AND METHODS

Purification of Flp:
Wild-type Flp and Flp variants were partially purified essentially as described by Prasad et al. (20). Strand cleavage and strand transfer assays were carried out with these preparations. Some reactions were done with 90 to 95% pure proteins obtained by an affinity purification protocol (18). Assays with affinity-pure proteins yielded the same results as those done with the less pure proteins. Fusion proteins composed of Flp (or a Flp variant) and glutathione S-transferase (GST) were purified in accordance with the procedure detailed by Yang and Jayaram (24). Protein concentrations were estimated by comparing densitometric scans of gel-fractionated aliquots stained with Coomassie brilliant blue to similar scans done with bovine serum albumin as the standard. These estimates were relatively crude and were only accurate within a factor of 2 or so.

Synthetic recombination sites:
Oligodeoxynucleotides for construction of full sites were synthesized in an Applied Biosystems 380A DNA synthesizer by using phosphoramidite chemistry (4). Normally, 10 to 20 pmol of each of the two appropriate oligodeoxynucleotide pairs was mixed in TE (10 mM Tris-HCl [pH 7.8] at 23 deg C, 1 mM EDTA [pH 8.0]),

intermediate which, following branch migration, is resolved into recombinants by the second pair of cleavage-joining reactions. The Int family recombinases use an active-site tyrosine as the nucleophile to attack the scissile phosphodiester during the strand breakage step. In Flp, this tyrosine residue is Tyr-343 (9). The active-site tyrosine is one of the invariant tetrad residues of the Int family (1, 3). The other three invariant residues are two arginines and a histidine (the RHR triad; Arg-191, His-305, and Arg-308 in Flp). The strand cleavage reaction results in covalent attachment of the recombinase to the 3(prm1) phosphate of DNA and exposure of a 5(prm1)-hydroxyl group at the nick. Strand joining in the recombinant mode is then effected via nucleophilic attack, by the 5(prm1)-hydroxyl group from the nicked strand of one DNA substrate, on the 3(prm1)-phosphotyrosyl bond formed within the partner substrate.

The Int family recombinases exist in solution as monomers and bind to DNA as monomers. Four recombinase monomers must act cooperatively to accomplish one round of recombination. Two concerted break exchanges must be made at one end of the strand exchange region (spacer) to form the Holliday junction. The process then needs to be repeated at the other end of the spacer to resolve the junction into mature recombinants. How does an Int family recombinase coordinate the breakage-joining events within the two DNA substrates taking part in recombination? Does the recombinase have a built-in mechanism by which it avoids abortive partial reactions within an incompletely assembled reaction complex?

The active-site configuration of two Int family members, Flp and the Zygosaccharomyces rouxii recombinase R, inferred from recombination reactions containing half-site substrates and step arrest mutants of the recombinases suggests potential solutions to the problems addressed above (6, 7, 13, 17, 24). A monomer of Flp or R harbors a partial active site; a complete active site is assembled by contribution of residues from more than one recombinase monomer. In the shared active site, three of the invariant Int family residues, the RHR triad (Arg-191, His-305, and Arg-308 in Flp; Arg-207, His-317, and Arg-320 in R), are derived from one monomer; the active-site tyrosine (Tyr-343 in Flp and Tyr-358 in R) is provided by a second monomer (6, 17). Assembly of a functional active site from partial active sites neatly accommodates the observation that in reactions with half sites, an Flp or an R monomer does not cleave the substrate to which it is bound but rather cleaves the substrate bound by a second recombinase monomer (trans DNA cleavage; 6). The partial active site, together with the trans mode of DNA cleavage, suggests possible mechanisms for postponing the chemistry of recombination until the complex is fully organized, for simultaneously assembling two active sites for coordinated strand cleavages, and for coupling the cleavage reaction with the conformational switch required for strand joining between partner substrates.

MOLECULAR AND CELLULAR BIOLOGY 1994

Title: Active-Site Assembly and Mode of DNA Cleavage by Flp Recombinase during Full-Site Recombination

Vol/Issue: 14 (1994)/11. (November)

Pages: 7492-7498

Author: ILSON WHANG, JEHEE LEE, MAKKUNI JAYARAM

Affiliation: Department of Microbiology, University of Texas at Austin, Austin, Texas 78712

Received: Received 10 June 1994/Returned for modification 4 August 1994/Accepted 12 August 1994

Abstract: A combination of half-site substrates and step arrest mutants of Flp, a site-specific recombinase of the integrase family, had earlier revealed the following features of the half-site recombination reaction. (i) The Flp active site is assembled by sharing of catalytic residues from at least two monomers of the protein. (ii) A Flp monomer does not cleave the half site to which it is bound (DNA cleavage in cis); rather, it cleaves a half site bound by a second Flp monomer (DNA cleavage in trans). For the (lambda) integrase (Int protein), the prototype member of the Int family, catalytic complementation between two active-site mutants has been observed in reactions with a suicide attL substrate. By analogy with Flp, this observation is strongly suggestive of a shared active site and of trans DNA cleavage. However, reactions with linear suicide attB substrates and synthetic Holliday junctions are more compatible with cis than with trans DNA cleavage. These Int results either argue against a common mode of active-site assembly within the Int family or challenge the validity of Flp half sites as mimics of the normal full-site substrates. We devised a strategy to assay catalytic complementation between Flp monomers in full sites. We found that the full-site reaction follows the shared active-site paradigm and the trans mode of DNA cleavage. These results suggest that within the Int family, a unitary chemical mechanism of recombination is achieved by more than one mode of physical interaction among the recombinase monomers.

Text:

The Flp protein of Saccharomyces cerevisiae is a conservative, site-specific DNA recombinase that belongs to the Int ((lambda) integrase) family of recombinases (1, 3). Members of this family execute recombination in two sequential steps. The first pair of strand cleavage-joining reactions produces a Holliday

172190 ----------

Definition **Yeast (S.cerevisiae) 2 micron circle plasmid, complete genome**

GenBank Name: YSCPLASM, Accession: J01347

NCBI Seq ID: 172190

Organism Saccharomyces cerevisiae

Comment [8] sites; mRNA CAP sites and poly-adenylation sites. [9] sites; FLP binding.
[7] sites; FLP cleavage.
[11] sites; FLP-mediated recombination crossover site. Draft entry and clean copy sequence for [5] kindly provided by J.Senecoff, 24-JAN-1986.
Yeast 2 micron plasmid contains two 599 bp inverted repeats separated by a large unique (UL) and a small unique (US) region. During recombination the UL and US regions invert producing two sequence forms that differ in the orientation of one unique region relative to the other. The A form is presented below. FLP is the only 2-micron circle-encoded protein needed for specific site recombination between the IRs of 2-micron circle. The minimal size of the recombination site required for efficient FLP recombinase-catalyzed recombination in vitro is no more than 28 bp, which includes parts of two 13 bp inverted repeats (positions 690-702 and 711-723) and all of an 8 bp spacer (703-710) [5]. The FLP recombinase cleaves the DNA at the boundaries of the spacer and becomes covalently linked to the spacer DNA [5],[9]. The efficiency of the recombination is reduced if the spacer in a recombinant site is increased or decreased by 1 bp, while the spacer in the second site is unaltered [5]. Recombination between two sites with identical 1-base pair additions or deletions is relatively unaffected, suggesting that pairing of sequences in the spacer regions is important in FLP-promoted recombination events [5]. The sequence asymmetry utilized by the recombinase to determine the orientation of the site is located uniquely within the spacer region. Another 13 bp direct repeat, is found at positions 676-688 [5]. FLP-mediated recombination involving two FLP sites that are inverted with respect to each other results in inversion of the DNA sequences between the sites [4]. If the participating recombination sites are in direct orientation, FLP promotes only the excision of the intervening DNA sequences [4]. The Rep 1 and Rep proteins are involved plasmid partitioning and protein stability.
A start codon in phase with the Rep1 coding region is located at positions 1966-1964. Two CAP sites for Rep1 mRNA are located beyond the 'atg' codon (position 2008) at positions 2004 and 2005. Complete source information:
Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71 [1], p82-6B [2], CV20 [3], pMMD2 [4], pGP20 [5], pJFS166 [10].

Updated Jul 31, 1992

Coding region	172190: 5570..6318 172190: 1..523
Coding region	172190: c2008..887
Coding region	172190: 2271..2816
Coding region	172190: c5198..4308

Sequence 6318 nt, circular ds genomic

```
   1 gaattctgaa ccagtcctaa aacgagtaaa taggaccggc aattcttcaa
   5 gcaataaaca ggaataccaa ttattaaaag ataacttagt cagatcgtac
 101 aataaagctt tgaagaaaaa tgcgccttat tcaatctttg ctataaaaaa
 151 tggcccaaaa tctcacattg gaagacattt gatgacctca tttctttcaa
 201 tgaagggcct aacggagttg actaatgttg tgggaaattg gagcgataag

 251 cgtgcttctg ccgtggccag gacaacgtat actcatcaga taacagcaat
 301 acctgatcac tacttcgcac tagtttctcg gtactatgca tatgatccaa
 351 tatcaaagga aatgatagca ttgaaggatg agactaatcc aattgaggag
 401 tggcagcata tagaacagct aaagggtagt gctgaaggaa gcatacgata
 451 ccccgcatgg aatgggataa tatcacagga ggtactagac taccttttcat

 501 cctacataaa tagacgcata taagtacgca tttaagcata aacacgcact
 551 atgccgttct tctcatgtat atatatatac aggcaacacg cagatatagg
 601 tgcgacgtga acagtgagct gtatgtgcgc agctcgcgtt gcatttttcgg
 651 aagcgctcgt tttcggaaac gctttgaagt tcctattccg aagttcctat
 701 tctctagaaa gtataggaac ttcagagcgc ttttgaaaac caaaagcgct

 751 ctgaagacgc actttcaaaa aaccaaaaac gcaccggact gtaacgagct
 801 actaaaatat tgcgaatacc gcttccacaa acattgctca aaagtatctc
 851 tttgctatat atctctgtgc tatatcccta tataacctac ccatccacct
 901 ttcgctcctt gaacttgcat ctaaactcga cctctacatc aacaggcttc
 951 caatgctctt caaattttac tgtcaagtag acccatacgg ctgtaatatg

1001 ctgctcttca taatgtaagc ttatctttat cgaatcgtgt gaaaaactac
1051 taccgcgata aacctttacg gttccctgag attgaattag ttcctttagt
1101 atatgataca agacactttt gaactttgta cgacgaattt tgaggttcgc
1151 catcctctgg ctatttccaa ttatcctgtc ggctattatc tccgcctcag
1201 tttgatcttc cgcttcagac tgccattttt cacataatga atctatttca

1251 ccccacaatc cttcatccgc ctccgcatct tgttccgtta aactattgac
1301 ttcatgttgt acattgttta gttcacgaga agggtcctct tcaggcggta
1351 gctcctgatc tcctatatga cctttatcct gttctctttc cacaaactta
1401 gaaatgtatt catgaattat ggagcaccta ataacattct tcaaggcgga
1451 gaagtttggg ccagatgccc aatatgcttg acatgaaaac gtgagaatga

1501 atttagtatt attgtgatat tctgaggcaa ttttattata atctcgaaga
1551 taagagaaga atgcagtgac ctttgtattg acaaatggag attccatgta
1601 tctaaaaaat acgcctttag gccttctgat acccttttccc ctgcggttta
1651 gcgtgccttt tacattaata tctaaaccct ctccgatggt ggcctttaac
```

1701 tgactaataa atgcaaccga tataaactgt gataattctg ggtgatttat

1751 gattcgatcg acaattgtat tgtacactag tgcaggatca ggccaatcca
1801 gttctttttc aattaccggt gtgtcgtctg tattcagtac atgtccaaca
1851 aatgcaaatg ctaacgtttt gtatttctta taattgtcag gaactggaaa
1901 agtccccctt gtcgtctcga ttacacacct actttcatcg tacaccatag
1951 gttggaagtg ctgcataata cattgcttaa tacaagcaag cagtctctcg

2001 ccattcatat ttcagttatt ttccattaca gctgatgtca ttgtatatca
2051 gcgctgtaaa aatctatctg ttacagaagg ttttcgcggt ttttataaac
2101 aaaactttcg ttacgaaatc gagcaatcac cccagctgcg tatttggaaa
2151 ttcgggaaaa agtagagcaa cgcgagttgc atttttttaca ccataatgca
2201 tgattaactt cgagaaggga ttaaggctaa tttcactagt atgtttcaaa

2251 aacctcaatc tgtccattga atgccttata aaacagctat agattgcata
2301 gaagagttag ctactcaatg ctttttgtca aagcttactg atgatgatgt
2351 gtctactttc aggcgggtct gtagtaagga gaatgacatt ataaagctgg
2401 cacttagaat tccacggact atagactata ctagtatact ccgtctactg
2451 tacgatacac ttccgctcag gtccttgtcc tttaacgagg ccttaccact

2501 cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat
2551 tctatcttcg cgatgtagta aaactagcta gaccgagaaa gagactagaa
2601 atgcaaaagg cacttctaca atggctgcca tcattattat ccgatgtgac
2651 gctgcagctt ctcaatgata ttcgaatacg ctttgaggag atacagccta
2701 atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca

2751 ttgaattttg aacatccgaa cctgggagtt ttccctgaaa cagatagtat
2801 atttgaacct gtataataat atatagtcta gcgctttacg gaagacaatg
2851 tatgtatttc ggttcctgga gaaactattg catctattgc ataggtaatc
2901 ttgcacgtcg catccccggt tcatttttctg cgtttccatc ttgcacttca
2951 atagcatatc tttgttaacg aagcatctgt gcttcatttt gtagaacaaa

3001 aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt
3051 tttacagaac agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat
3101 ctgtgcttca tttttgtaaa acaaaaatgc aacgcgagag cgctaatttt
3151 tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga
3201 gagcgctatt ttaccaacaa agaatctata cttctttttt gttctacaaa

3251 aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt
3301 ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac
3351 tgtaggtccg ttaaggttag aagaaggcta ctttggtgtc tattttctct
3401 tccataaaaa aagcctgact ccacttcccg cgtttactga ttactagcga
3451 agctgcgggt gcatttttttc aagataaagg catccccgat tatattctat

3501 accgatgtgg attgcgcata ctttgtgaac agaaagtgat agcgttgatg
3551 attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta
3601 tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact
3651 ctatgaatag ttcttactac aatttttttg tctaaagagt aatactagag
3701 ataaacataa aaaatgtaga ggtcgagttt agatgcaagt tcaaggagcg

3751 aaaggtggat gggtaggtta tatagggata tagcacagag atatatagca

3801 aagagatact tttgagcaat gtttgtggaa gcggtattcg caatatttta
3851 gtagctcgtt acagtccggt gcgtttttgg ttttttgaaa gtgcgtcttc
3901 agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga
3951 gaataggaac ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc

4001 ttccgaaaat gcaacgcgag ctgcgcacat acagctcact gttcacgtcg
4051 cacctatatc tgcgtgttgc ctgtatatat atatacatga gaagaacggc
4101 atagtgcgtg tttatgctta aatgcgtact tatatgcgtc tatttatgta
4151 ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg
4201 ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc

4251 cactcctcaa ttggattagt ctcatccttc aatgctatca tttcctttga
4301 tattggatca taccctagaa gtattacgtg attttctgcc ccttaccctc
4351 gttgctactc tccttttttt cgtgggaacc gctttagggc cctcagtgat
4401 ggtgttttgt aatttatatg ctcctcttgc atttgtgtct ctacttcttg
4451 ttcgcctgga gggaacttct tcatttgtat tagcatggtt cacttcagtc

4501 cttccttcca actcactctt tttttgctgt aaacgattct ctgccgccag
4551 ttcattgaaa ctattgaata tatcctttag agattccggg atgaataaat
4601 cacctattaa agcagcttga cgatctggtg gaactaaagt aagcaattgg
4651 gtaacgacgc ttacgagctt cataacatct tcttccgttg gagctggtgg
4701 gactaataac tgtgtacaat ccatttttct catgagcatt tcggtagctc

4751 tcttcttgtc tttctcgggc aatcttccta ttattatagc aatagatttg
4801 tatagttgct ttctattgtc taacagcttg ttattctgta gcatcaaatc
4851 tatggcagcc tgacttgctt cttgtgaaga gagcatacca tttccaatcg
4901 aatcaaacct ttccttaacc atcttcgcag caggcaaaat tacctcagca
4951 ctggagtcag aagatacgct ggaatcttct gcgctagaat caagaccata

5001 cggcctaccg gttgtgagag attccatggg ccttatgaca tatcctggaa
5051 agagtagctc atcagactta cgtttactct ctatatcaat atctacatca
5101 ggagcaatca tttcaataaa cagccgacat acatcccaga cgctataagc
5151 tgtacgtgct tttaccgtca gattcttggc tgtttcaatg tcgtccattt
5201 tggttttctt ttaccagtat tgttcgtttg ataatgtatt cttgcttatt

5251 acattataaa atctgtgcag atcacatgtc aaaacaactt tttatcacaa
5301 gatagtaccg caaaacgaac ctgcgggccg tctaaaaatt aaggaaaagc
5351 agcaaaggtg catttttaaa atatgaaatg aagataccgc agtaccaatt
5401 attttcgcag tacaaataat gcgcggccgg tgcatttttc gaaagaacgc
5451 gagacaaaca ggacaattaa agttagtttt tcgagttagc gtgtttgaat

5501 actgcaagat acaagataaa tagagtagtt gaaactagat atcaattgca
5551 cacaagatcg gcgctaagca tgccacaatt tggtatatta tgtaaaacac
5601 cacctaaggt gcttgttcgt cagtttgtgg aaaggtttga aagaccttca
5651 ggtgagaaaa tagcattatg tgctgctgaa ctaacctatt tatgttggat
5701 gattacacat aacggaacag caatcaagag agccacattc atgagctata

5751 atactatcat aagcaattcg ctgagtttcg atattgtcaa taaatcactc
5801 cagtttaaat acaagacgca aaaagcaaca attctggaag cctcattaaa
5851 gaaattgatt cctgcttggg aatttacaat tattccttac tatggacaaa
5901 aacatcaatc tgatatcact gatattgtaa gtagtttgca attacagttc

5951 gaatcatcgg aagaagcaga taagggaaat agccacagta aaaaaatgct

6001 taaagcactt ctaagtgagg gtgaaagcat ctgggagatc actgagaaaa
6051 tactaaattc gtttgagtat acttcgagat ttacaaaaac aaaaacttta
6101 taccaattcc tcttcctagc tactttcatc aattgtggaa gattcagcga
6151 tattaagaac gttgatccga aatcatttaa attagtccaa aataagtatc
6201 tgggagtaat aatccagtgt ttagtgacag agacaaagac aagcgttagt

6251 aggcacatat acttctttag cgcaaggggt aggatcgatc cacttgtata
6301 tttggatgaa tttttgag

08/486,409

```
PLEASE ENTER HOST PORT ID:x
LOGINID:d184csl
PASSWORD:
TERMINAL (ENTER 1, 2, 3, 4, OR ?):2e
06 JUN 92 14:03:30          U.S. Patent & Trademark Office          P0001
=> set pagelength scroll
SET COMMAND COMPLETED

=> s flp(5a)recombinas?
           151 FLP
             6 RECOMBINAS?
L1           1 FLP(5A)RECOMBINAS?

=> d

1.  4,997,757, Mar. 5, 1991, Process for detecting potential carcinogens;
Robert H. Schiestl, 435/172.1, 6, 29, 172.3; 935/76, 78, 79, 84 [IMAGE
AVAILABLE]

=>


 Alt-Z FOR HELP| VT102    | FDX |  2400 N81 | LOG CLOSED | PRINT OFF | ON-LINE
d his

     (FILE 'USPAT' ENTERED AT 14:03:27 ON 06 JUN 92)
                SET PAGELENGTH SCROLL
L1           1 S FLP(5A)RECOMBINAS?
L2           1 S L1 AND MAMMAL?

     FILE 'JPOABS' ENTERED AT 14:06:27 ON 06 JUN 92
L3           0 S L1

     FILE 'USPAT' ENTERED AT 14:06:43 ON 06 JUN 92

=> d
'L3' HAS NO ANSWERS
L3           0 SEA FLP(5A)RECOMBINAS?

=> d l1

1.  4,997,757, Mar. 5, 1991, Process for detecting potential carcinogens;
Robert H. Schiestl, 435/172.1, 6, 29, 172.3; 935/76, 78, 79, 84 [IMAGE
AVAILABLE]

=> file jpoabs
FILE 'JPOABS' ENTERED AT 14:07:11 ON 06 JUN 92

  * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  *    J A P A N E S E   P A T E N T   A B S T R A C T S        *
  *                                                             *
  * CURRENTLY, DATA IS LOADED THROUGH THE ABSTRACT PUBLICATION  *
  * DATE OF AUGUST 30, 1991.                                    *
  * THE LATEST GROUPS RECEIVED ARE: C0862 E1105, M1150 & P1245. *
  * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *


=> s l1
            21 FLP
             0 RECOMBINAS?
L4           0 FLP(5A)RECOMBINAS?
```

APS

~~07/660,252~~

t2/3/1-130

2/3/1 (Item 1 from file: 155)
08128396 92266396
DNA cleavage in trans by the active site tyrosine during Flp recombination: switching protein partners before exchanging strands.
Chen JW; Lee J; Jayaram M
Department of Microbiology, University of Texas, Austin 78712.
Cell (UNITED STATES) May 15 1992, 69 (4) p647-58, ISSN 0092-8674
Journal Code: CQ4
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/2 (Item 2 from file: 155)
08080444 92218444
Reactions between half- and full-FLP recombination target sites. A model system for analyzing early steps in FLP protein-mediated site-specific recombination.
Qian XH; Inman RB; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison 53706.
J Biol Chem (UNITED STATES) Apr 15 1992, 267 (11) p7794-805, ISSN 0021-9258 Journal Code: HIV
Contract/Grant No.: GM-32335; GM-14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/3 (Item 3 from file: 155)
07913378 92051378
FLP-mediated recombination in the vector mosquito, Aedes aegypti.
Morris AC; Schaub TL; James AA
Department of Molecular Biology & Biochemistry, University of California, Irvine 92717.
Nucleic Acids Res (ENGLAND) Nov 11 1991, 19 (21) p5895-900, ISSN 0305-1048 Journal Code: O8L
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/4 (Item 4 from file: 155)
07823652 91342652
Synapsis, strand scission, and strand exchange induced by the FLP recombinase: analysis with half-FRT sites.
Amin A; Roca H; Luetke K; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
Mol Cell Biol Sep 1991, 11 (9) p4497-508, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/5 (Item 5 from file: 155)
07777737 91296737
Domain of a yeast site-specific recombinase (Flp) that recognizes its target site.

Chen JW; Evans BR; Yang SH; Teplow DB; Jayaram M
Department of Microbiology, University of Texas, Austin 78712.
Proc Natl Acad Sci U S A Jul 15 1991, 88 (14) p5944-8, ISSN 0027-8424 Journal Code: PV3
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/6 (Item 6 from file: 155)
07731454 91250454
Identification of the DNA-binding domain of the FLP recombinase.
Pan H; Clary D; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Biol Chem Jun 15 1991, 266 (17) p11347-54, ISSN 0021-9258 Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/7 (Item 7 from file: 155)
07687992 91206992
Integration specificity of retrotransposons and retroviruses.
Sandmeyer SB; Hansen LJ; Chalker DL
Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine 92717.
Annu Rev Genet 1990, 24 p491-518, ISSN 0066-4197 Journal Code: 6DP
Contract/Grant No.: GM33281
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, ACADEMIC

2/3/8 (Item 8 from file: 155)
07668658 91187658
A bacterial model system for chromosomal targeting.
Huang LC; Wood EA; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
Nucleic Acids Res Feb 11 1991, 19 (3) p443-8, ISSN 0305-1048 Journal Code: O8L
Contract/Grant No.: GM37835
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/9 (Item 9 from file: 155)
07645850 91164850
Recombinase-mediated gene activation and site-specific integration in mammalian cells.
O'Gorman S; Fox DT; Wahl GM
Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037.
Science Mar 15 1991, 251 (4999) p1351-5, ISSN 0036-8075 Journal Code: UJ7
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/10 (Item 10 from file: 155)
07643634 91162634

Tyr60 variants of Flp recombinase generate conformationally altered protein-DNA complexes. Differential activity in full-site and half-site recombinations.
Chen JW; Evans BR; Zheng L; Jayaram M
Department of Microbiology, University of Texas at Austin, Austin 78712.
J Mol Biol Mar 5 1991, 218 (1) p107-18, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/11 (Item 11 from file: 155)
07554393 91073393
FLP protein of 2 mu circle plasmid of yeast induces multiple bends in the FLP recognition target site.
Schwartz CJ; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Nov 20 1990, 216 (2) p289-98, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/12 (Item 12 from file: 155)
07553382 91072382
Protein-based asymmetry and protein-protein interactions in FLP recombinase-mediated site-specific recombination.
Qian XH; Inman RB; Cox MM
Program in Cell and Molecular Biology, College of Agricultural and Life Sciences, University of Wisconsin, Madison 53706.
J Biol Chem Dec 15 1990, 265 (35) p21779-88, ISSN 0021-9258
Journal Code: HIV
Contract/Grant No.: GM 37835; GM 14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/13 (Item 13 from file: 155)
07490349 91009349
Identification of the active site tyrosine of Flp recombinase. Possible relevance of its location to the mechanism of recombination [published erratum appears in J Biol Chem 1991 Apr 15;266(11):7312]
Evans BR; Chen JW; Parsons RL; Bauer TK; Teplow DB; Jayaram M
Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037.
J Biol Chem Oct 25 1990, 265 (30) p18504-10, ISSN 0021-9258
Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/14 (Item 14 from file: 155)
07410836 90317836
Synaptic intermediates promoted by the FLP recombinase.
Amin AA; Beatty LG; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Jul 5 1990, 214 (1) p55-72, ISSN 0022-2836
Journal Code: J6V

Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/15 (Item 15 from file: 155)
07263960 90170960
Functional analysis of Arg-308 mutants of Flp recombinase. Possible role of Arg-308 in coupling substrate binding to catalysis.
Parsons RL; Evans BR; Zheng L; Jayaram M
Research Institute of Scripps Clinic, La Jolla, California 92037.
J Biol Chem Mar 15 1990, 265 (8) p4527-33, ISSN 0021-9258 Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/16 (Item 16 from file: 155)
07229522 90136522
Use of site-specific recombination to regenerate selectable markers.
Cregg JM; Madden KR
Salk Institute Biotechnology/Industrial Associates, Inc., La Jolla, CA 92037.
Mol Gen Genet Oct 1989, 219 (1-2) p320-3, ISSN 0026-8925 Journal Code: NGP
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/17 (Item 17 from file: 155)
07190832 90097832
Characterization of Holliday structures in FLP protein-promoted site-specific recombination.
Meyer-Leon L; Inman RB; Cox MM
Program in Cellular and Molecular Biology, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706-1569.
Mol Cell Biol Jan 1990, 10 (1) p235-42, ISSN 0270-7306 Journal Code: NGY
Contract/Grant No.: GM37835; GM14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/18 (Item 18 from file: 155)
07123422 90030422
The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome.
Golic KG; Lindquist S
Howard Hughes Medical Institute, Department of Molecular Genetics and Cell Biology, University of Chicago, Illinois 60637.
Cell Nov 3 1989, 59 (3) p499-509, ISSN 0092-8674 Journal Code: CQ4
Contract/Grant No.: GM 25874
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/19 (Item 19 from file: 155)
07011744 89313744
Synthesis of an enzymatically active FLP recombinase in vitro: search for a DNA-binding domain.

Amin AA; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
Mol Cell Biol May 1989, 9 (5) p1987-95, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/20 (Item 20 from file: 155)
07002130 89304130
FLP-FRT mediated intrachromosomal recombination on a tandemly duplicated YEp integrant at the ILV2 locus of chromosome XIII in Saccharomyces cerevisiae.
Rank GH; Arndt GM; Xiao W
Department of Biology, University of Saskatchewan, Saskatoon, Canada.
Curr Genet Feb 1989, 15 (2) p107-12, ISSN 0172-8083 Journal Code: CUG
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/21 (Item 21 from file: 155)
06876684 89178684
FLP recombinase of the 2 microns circle plasmid of Saccharomyces cerevisiae bends its DNA target. Isolation of FLP mutants defective in DNA bending.
Schwartz CJ; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Feb 20 1989, 205 (4) p647-58, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/22 (Item 22 from file: 155)
06825220 89127220
Holliday intermediates and reaction by-products in FLP protein-promoted site-specific recombination.
Meyer-Leon L; Huang LC; Umlauf SW; Cox MM; Inman RB
Department of Biochemistry, College of Agriculture and Life Sciences, University of Wisconsin-Madison 53706-1569.
Mol Cell Biol Sep 1988, 8 (9) p3784-96, ISSN 0270-7306
Journal Code: NGY
Contract/Grant No.: GM37835; GM14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/23 (Item 23 from file: 155)
06823587 89125587
The mechanism of loading of the FLP recombinase onto its DNA target sequence.
Beatty LG; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Nov 20 1988, 204 (2) p283-94, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/24 (Item 24 from file: 155)
06794920 89096920
Step-arrest mutants of FLP recombinase: implications for the catalytic mechanism of DNA recombination.
Parsons RL; Prasad PV; Harshey RM; Jayaram M
Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037.
Mol Cell Biol Aug 1988, 8 (8) p3303-10, ISSN 0270-7306 Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/25 (Item 25 from file: 155)
06761437 89063437
High frequency FLP-independent homologous DNA recombination of 2 mu plasmid in the yeast Saccharomyces cerevisiae.
Bruschi CV; Howe GA
Department of Microbiology and Immunology, School of Medicine, East Carolina University, Greenville, NC 27858-4354.
Curr Genet Sep 1988, 14 (3) p191-9, ISSN 0172-8083 Journal Code: CUG
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/26 (Item 26 from file: 155)
06740094 89042094
Holliday junctions in FLP recombination: resolution by step-arrest mutants of FLP protein.
Jayaram M; Crain KL; Parsons RL; Harshey RM
Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, CA 92037.
Proc Natl Acad Sci U S A Nov 1988, 85 (21) p7902-6, ISSN 0027-8424 Journal Code: PV3
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/27 (Item 27 from file: 155)
06703077 89005077
The functional significance of DNA sequence structure in a site-specific genetic recombination reaction.
Umlauf SW; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
EMBO J Jun 1988, 7 (6) p1845-52, ISSN 0261-4189 Journal Code: EMB
Contract/Grant No.: GM37835; AI00599; GM07215
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/28 (Item 28 from file: 155)
06687975 88332975
DNA recognition by the FLP recombinase of the yeast 2 mu plasmid. A mutational analysis of the FLP binding site.
Senecoff JF; Rossmeissl PJ; Cox MM

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
J Mol Biol May 20 1988, 201 (2) p405-21, ISSN 0022-2836
Journal Code: J6V
Contract/Grant No.: GM37835; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/29 (Item 29 from file: 155)
06643050 88288050
Nucleotide sequencing and expression of the fadL gene involved in long-chain fatty acid transport in Escherichia coli.
Said B; Ghosn CR; Vu L; Nunn WD
Department of Molecular Biology and Biochemistry, University of California, Irvine 92717.
Mol Microbiol May 1988, 2 (3) p363-70, ISSN 0950-382X
Journal Code: MOM
Contract/Grant No.: GM 22466-11
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/30 (Item 30 from file: 155)
06618001 88263001
FLP recombinase is an enzyme.
Gates CA; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
Proc Natl Acad Sci U S A Jul 1988, 85 (13) p4628-32, ISSN 0027-8424
Journal Code: PV3
Contract/Grant No.: GM37835; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/31 (Item 31 from file: 155)
06567126 88212126
Mutations that improve the binding of yeast FLP recombinase to its substrate.
Lebreton B; Prasad PV; Jayaram M; Youderian P
Department of Biological Sciences, University of Southern California, Los Angeles 90089-1481.
Genetics Mar 1988, 118 (3) p393-400, ISSN 0016-6731 Journal Code: FNH
Contract/Grant No.: GM34982; GM35654
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/32 (Item 32 from file: 155)
06521666 88166666
Antagonistic controls regulate copy number of the yeast 2 mu plasmid.
Murray JA; Scarpa M; Rossi N; Cesareni G
EMBL, Heidelberg, FRG.
EMBO J Dec 20 1987, 6 (13) p4205-12, ISSN 0261-4189 Journal Code: EMB
Languages: ENGLISH

Document type: JOURNAL ARTICLE

2/3/33 (Item 33 from file: 155)
06506025 88151025
Autoregulation of 2 micron circle gene expression provides a model for maintenance of stable plasmid copy levels.
Som T; Armstrong KA; Volkert FC; Broach JR
Department of Molecular Biology, Princeton University, New Jersey 08544.
Cell Jan 15 1988, 52 (1) p27-37, ISSN 0092-8674 Journal Code: CQ4
Contract/Grant No.: GM34596
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/34 (Item 34 from file: 155)
06342913 87316913
Purification of the FLP site-specific recombinase by affinity chromatography and re-examination of basic properties of the system.
Meyer-Leon L; Gates CA; Attwood JM; Wood EA; Cox MM
Nucleic Acids Res Aug 25 1987, 15 (16) p6469-88, ISSN 0305-1048 Journal Code: O8L
Contract/Grant No.: GM32335; GM37835; AI00599; +
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/35 (Item 35 from file: 155)
06280212 87254212
Isolation of intermediates in the binding of the FLP recombinase of the yeast plasmid 2-micron circle to its target sequence.
Andrews BJ; Beatty LG; Sadowski PD
J Mol Biol Jan 20 1987, 193 (2) p345-58, ISSN 0022-2836 Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/36 (Item 36 from file: 155)
06274060 87248060
Rapid localization and characterization of random mutations within the 2 micron circle site-specific recombinase: a general strategy for analysis of protein function [published erratum appears in Gene 1987;57(1):149]
Govind NS; Jayaram M
Gene 1987, 51 (1) p31-41, ISSN 0378-1119 Journal Code: FOP
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/37 (Item 37 from file: 155)
06210407 87184407
Site-specific recombination of the yeast plasmid two-micron circle: intermediates in the binding process.
Andrews BJ; Beatty LG; Sadowski PD
Basic Life Sci 1986, 40 p407-24, ISSN 0090-5542 Journal Code: 9K0
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/38 (Item 38 from file: 155)

06210406 87184406
Site-specific recombination promoted in vitro by the FLP protein of the yeast two-micron plasmid.
Senecoff JF; Bruckner RC; Meyer-Leon L; Gates CA; Wood E; Umlauf SW; Attwood JM; Cox MM
Basic Life Sci 1986, 40 p397-405, ISSN 0090-5542 Journal Code: 9K0
Contract/Grant No.: GM32335; 5-T32 GM07215; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/39 (Item 39 from file: 155)
06210404 87184404
Survival strategies of the yeast plasmid two-micron circle.
Volkert FC; Wu LC; Fisher PA; Broach JR
Basic Life Sci 1986, 40 p375-96, ISSN 0090-5542 Journal Code: 9K0
Contract/Grant No.: GM34596; GM33132
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/40 (Item 40 from file: 155)
06201639 87175639
Mutations in the 2-microns circle site-specific recombinase that abolish recombination without affecting substrate recognition [published erratum appears in Proc Natl Acad Sci U S A 1988 Mar;85(5):1497]
Prasad PV; Young LJ; Jayaram M
Proc Natl Acad Sci U S A Apr 1987, 84 (8) p2189-93, ISSN 0027-8424 Journal Code: PV3
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/41 (Item 41 from file: 155)
06167165 87141165
Association of reciprocal exchange with gene conversion between the repeated segments of 2-micron circle.
Jayaram M
J Mol Biol Oct 5 1986, 191 (3) p341-54, ISSN 0022-2836 Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/42 (Item 42 from file: 155)
06115790 87089790
Substrate recognition by the 2 micron circle site-specific recombinase: effect of mutations within the symmetry elements of the minimal substrate.
Prasad PV; Horensky D; Young LJ; Jayaram M
Mol Cell Biol Dec 1986, 6 (12) p4329-34, ISSN 0270-7306 Journal Code: NGY
Contract/Grant No.: GM 35654-01
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/43 (Item 43 from file: 155)
06115725 87089725
Mating type-like conversion promoted by the 2 micrograms circle

site-specific recombinase: implications for the double-strand-gap repair model.
Jayaram M
Mol Cell Biol Nov 1986, 6 (11) p3831-7, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/44 (Item 44 from file: 155)
06115667 87089667
Identification of the crossover site during FLP-mediated recombination in the Saccharomyces cerevisiae plasmid 2 microns circle.
McLeod M; Craft S; Broach JR
Mol Cell Biol Oct 1986, 6 (10) p3357-67, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/45 (Item 45 from file: 155)
06090546 87064546
Interaction of the FLP recombinase of the Saccharomyces cerevisiae 2 micron plasmid with mutated target sequences.
Andrews BJ; McLeod M; Broach J; Sadowski PD
Mol Cell Biol Jul 1986, 6 (7) p2482-9, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/46 (Item 46 from file: 155)
06009798 86310798
The FLP recombinase of the Saccharomyces cerevisiae 2 microns plasmid attaches covalently to DNA via a phosphotyrosyl linkage.
Gronostajski RM; Sadowski PD
Mol Cell Biol Nov 1985, 5 (11) p3274-9, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/47 (Item 47 from file: 155)
06003314 86304314
Specific contacts between the FLP protein of the yeast 2-micron plasmid and its recombination site.
Bruckner RC; Cox MM
J Biol Chem Sep 5 1986, 261 (25) p11798-807, ISSN 0021-9258
Journal Code: HIV
Contract/Grant No.: GM32335; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/48 (Item 48 from file: 155)
05983659 86284659
Chromatin organization of the Saccharomyces cerevisiae 2 microns plasmid depends on plasmid-encoded products.
Veit BE; Fangman WL

Mol Cell Biol Sep 1985, 5 (9) p2190-6, ISSN 0270-7306
Journal Code: NGY
Contract/Grant No.: GM18926
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/49 (Item 49 from file: 155)
05980709 86281709
FLP site-specific recombinase of yeast 2-micron plasmid. Topological features of the reaction.
Beatty LG; Babineau-Clary D; Hogrefe C; Sadowski PD
J Mol Biol Apr 20 1986, 188 (4) p529-44, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/50 (Item 50 from file: 155)
05971102 86272102
Site-specific recombination promotes plasmid amplification in yeast.
Volkert FC; Broach JR
Cell Aug 15 1986, 46 (4) p541-50, ISSN 0092-8674 Journal Code: CQ4
Contract/Grant No.: GM-34596
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/51 (Item 51 from file: 155)
05958059 86259059
The minimal duplex DNA sequence required for site-specific recombination promoted by the FLP protein of yeast in vitro.
Proteau G; Sidenberg D; Sadowski P
Nucleic Acids Res Jun 25 1986, 14 (12) p4787-802, ISSN 0305-1048
Journal Code: O8L
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/52 (Item 52 from file: 155)
05931585 86232585
Sequence organization of the circular plasmid pKD1 from the yeast Kluyveromyces drosophilarum.
Chen XJ; Saliola M; Falcone C; Bianchi MM; Fukuhara H
Nucleic Acids Res Jun 11 1986, 14 (11) p4471-81, ISSN 0305-1048
Journal Code: O8L
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/53 (Item 53 from file: 155)
05923006 86224006
Directionality in FLP protein-promoted site-specific recombination is mediated by DNA-DNA pairing.
Senecoff JF; Cox MM
J Biol Chem Jun 5 1986, 261 (16) p7380-6, ISSN 0021-9258
Journal Code: HIV
Contract/Grant No.: GM32335; AI00599
Languages: ENGLISH

Document type: JOURNAL ARTICLE

2/3/54 (Item 54 from file: 155)
05919123 86220123
The integrase family of site-specific recombinases: regional similarities and global diversity.
Argos P; Landy A; Abremski K; Egan JB; Haggard-Ljungquist E; Hoess RH; Kahn ML; Kalionis B; Narayana SV; Pierson LS 3d; et al
EMBO J Feb 1986, 5 (2) p433-40, ISSN 0261-4189 Journal Code: EMB
Contract/Grant No.: AI 13544
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/55 (Item 55 from file: 155)
05810590 86111590
Site-specific recombinases: changing partners and doing the twist.
Sadowski P
J Bacteriol Feb 1986, 165 (2) p341-7, ISSN 0021-9193 Journal Code: HH3
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW

2/3/56 (Item 56 from file: 155)
05741647 86042647
The FLP recombinase of the yeast 2-micron plasmid: characterization of its recombination site.
Senecoff JF; Bruckner RC; Cox MM
Proc Natl Acad Sci U S A Nov 1985, 82 (21) p7270-4, ISSN 0027-8424 Journal Code: PV3
Contract/Grant No.: GM32335
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/57 (Item 57 from file: 155)
05707309 86008309
The FLP protein of the 2-micron plasmid of yeast. Inter- and intramolecular reactions.
Gronostajski RM; Sadowski PD
J Biol Chem Oct 5 1985, 260 (22) p12328-35, ISSN 0021-9258 Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/58 (Item 58 from file: 155)
05707308 86008308
Determination of DNA sequences essential for FLP-mediated recombination by a novel method.
Gronostajski RM; Sadowski PD
J Biol Chem Oct 5 1985, 260 (22) p12320-7, ISSN 0021-9258 Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/59 (Item 59 from file: 155)

05707307 86008307
The FLP protein of the 2-micron plasmid of yeast. Purification of the protein from Escherichia coli cells expressing the cloned FLP gene.
Babineau D; Vetter D; Andrews BJ; Gronostajski RM; Proteau GA; Beatty LG; Sadowski PD
J Biol Chem Oct 5 1985, 260 (22) p12313-9, ISSN 0021-9258
Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/60 (Item 60 from file: 155)
05560933 85176933
The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences.
Andrews BJ; Proteau GA; Beatty LG; Sadowski PD
Cell Apr 1985, 40 (4) p795-803, ISSN 0092-8674 Journal Code: CQ4
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/61 (Item 1 from file: 5)
8906509 BIOSIS Number: 42131509
AN ORDERED DISASSEMBLY OF COMPLEXES OF FLP RECOMBINASE AND FRT SITES FOLLOWING RECOMBINATION
WAITE L L; COX M M
DEP. BIOCHEM., UNIV. WISCONSIN, MADISON, WIS. 53706.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 67. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/62 (Item 2 from file: 5)
8906501 BIOSIS Number: 42131501
LIGATION ACTIVITY OF THE FLP RECOMBINASE
PAN G; SADOWSKI P D
DEP. MOLECULAR MED. GENETICS, UNIV. TORONTO, TORONTO, ONTARIO M5S 1A8, CAN.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 65. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/63 (Item 3 from file: 5)
8906498 BIOSIS Number: 42131498
HALF-SITE RECOMBINATIONS MEDIATED BY FLP RECOMBINASE FROM SACCHAROMYCES-CEREVISIAE
SERRE M-C; LEI-ZHENG; JAYARAM M
DEP. MICROBIOL., UNIV. TEXAS AUSTIN, AUSTIN, TEX. 78746.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 64. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/64 (Item 4 from file: 5)
8906492 BIOSIS Number: 42131492
FUNCTIONAL ANALYSES OF MUTANTS OF FLP AND R RECOMBINASE FROM YEAST
CHEN J-W; LEE J; EVANS B; SERRE M-C; ARAKI H; OSHIMA Y; JAYARAM M
DEP. MICROBIOL., UNIV. TEXAS AUSTIN, AUSTIN, TEX. 78712.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 62. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/65 (Item 5 from file: 5)
8197568 BIOSIS Number: 91118568
TYROSINE-60 VARIANTS OF FLP RECOMBINASE GENERATE CONFORMATIONALLY ALTERED PROTEIN DNA COMPLEXES DIFFERENTIAL ACTIVITY IN FULL-SITE AND HALF RECOMBINATIONS
CHEN J-W; EVANS B R; ZHENG L; JAYARAM M
DEP. MICRBOIOL., UNIV. TEXAS AUSTIN, AUSTIN, TEX. 78712, USA.
J MOL BIOL 218 (1). 1991. 107-118. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/66 (Item 6 from file: 5)
7103760 BIOSIS Number: 88026505
FLP-FRT MEDIATED INTRACHROMOSOMAL RECOMBINATION ON A TANDEMLY DUPLICATED YE-P INTEGRANT AT THE ILV2 LOCUS OF CHROMOSOME XIII IN SACCHAROMYCES-CEREVISIAE
RANK G H; ARNDT G M; XIAO W
DEP. BIOL., UNIV. SASKATCHEWAN, SASKATOON, SASKATCHEWAN, CANADA S7N 0W0.
CURR GENET 15 (2). 1989. 107-112. CODEN: CUGED
Full Journal Title: Current Genetics
Language: ENGLISH

2/3/67 (Item 7 from file: 5)
7043154 BIOSIS Number: 87103675
FLP RECOMBINASE OF THE 2 MUM CIRCLE PLASMID OF SACCHAROMYCES-CEREVISIAE BENDS ITS DNA TARGET ISOLATION OF FLP MUTANTS DEFECTIVE IN DNA BENDING
SCHWARTZ C J E; SADOWSKI P D
DEP. MED. GENETICS, UNIV. TORONTO, TORONTO, ONTARIO M5S 1A8, CAN.
J MOL BIOL 205 (4). 1989. 647-658. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/68 (Item 8 from file: 5)
6944460 BIOSIS Number: 87004981
HIGH FREQUENCY FLP-INDEPENDENT HOMOLOGOUS DNA RECOMBINATION OF 2 MICRON PLASMID IN THE YEAST SACCHAROMYCES-CEREVISIAE
BRUSCHI C V; HOWE G A
DEP. MICROBIOL. IMMUNOL., SCH. MED., EAST CAROLINA UNIV., GREENVILLE, N.C. 27858-4354, U.S.A.
CURR GENET 14 (3). 1988. 191-200. CODEN: CUGED
Full Journal Title: Current Genetics
Language: ENGLISH

2/3/69 (Item 9 from file: 5)
6892306 BIOSIS Number: 37086685
THE FLP RECOMBINASE STEP-ARREST MUTANTS AND INTERMEDIATES IN RECOMBINATION
JAYARAM M; PARSONS R; EVANS B
RES. INST. SCRIPPS CLIN., LA JOLLA, CALIF. 92037.
SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION HELD AT THE 18TH ANNUAL UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, STEAMBOAT SPRINGS, COLORADO, USA, MARCH 27-APRIL 3, 1989. J CELL BIOCHEM SUPPL 0 (13 PART D). 1989. 106. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/70 (Item 10 from file: 5)
6636107 BIOSIS Number: 86102658
AUTOREGULATION OF 2-MUM CIRCLE GENE EXPRESSION PROVIDES A MODEL FOR MAINTENANCE OF STABLE PLASMID COPY LEVELS
SOM T; ARMSTRONG K A; VOLKERT F C; BROACH J R
DEP. MOLECULAR BIOL., PRINCETON UNIV., PRINCETON, NEW JERSEY 08544.
CELL 52 (1). 1988. 27-38. CODEN: CELLB
Full Journal Title: Cell
Language: ENGLISH

2/3/71 (Item 11 from file: 5)
6624830 BIOSIS Number: 86091381
THE INT FAMILY OF SITE-SPECIFIC RECOMBINASES SOME THOUGHTS ON A GENERAL REACTION MECHANISM
JAYARAM M
DEP. MOL. BIOL., RES. INST. SCRIPPS CLINIC, 10666 NORTH TORREY PINES ROAD, LA JOLLA, CALIF. 92037, USA.
J GENET 67 (1). 1988. 29-36. CODEN: JOGNA
Full Journal Title: Journal of Genetics
Language: ENGLISH

2/3/72 (Item 12 from file: 5)
6571174 BIOSIS Number: 86037725
FLP RECOMBINASE INDUCTION OF THE BREAKAGE-FUSION-BRIDGE CYCLE AND GENE CONVERSION IN SACCHAROMYCES-CEREVISIAE
RANK G H; XIAO W; KOLENOVSKY A; ARNDT G
DEP. BIOL., UNIV. SASK., SASKATOON, SASK., CAN. S7N OWO.
CURR GENET 13 (4). 1988. 273-282. CODEN: CUGED
Full Journal Title: Current Genetics
Language: ENGLISH

2/3/73 (Item 13 from file: 5)
6150196 BIOSIS Number: 35015717
PURIFICATION OF FLP RECOMBINASE USING SEQUENCE-SPECIFIC DNA AFFINITY CHROMATOGRAPHY
GATES C A; MEYER-LEON L; ATTWOOD J M; WOOD E A; COX M M
DEP. BIOCHEM., UNIV. WIS.-MADISON, MADISON, WIS. 53706, USA.
BURGESS, R. (ED.). UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY NEW SERIES, VOL. 68. PROTEIN

PURIFICATION: MICRO TO MACRO; CETUS-UCLA SYMPOSIUM, FRISCO, COLORADO, USA, MARCH 29-APRIL 4, 1987. XVIII+510P. ALAN R. LISS, INC.: NEW YORK, NEW YORK, USA. ILLUS. ISBN 0-8451-2667-9. 0 (0). 1987. 197-206. CODEN: USMBD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/74 (Item 14 from file: 5)
5802738 BIOSIS Number: 83065045
SUBSTRATE RECOGNITION BY THE 2-MICROMETER CIRCLE SITE-SPECIFIC RECOMBINASE EFFECT OF MUTATIONS WITHIN THE SYMMETRY ELEMENTS OF THE MINIMAL SUBSTRATE
PRASAD P V; HORENSKY D; YOUNG L-J; JAYARAM M
DEP. MOL. BIOL., RES. INST. SCRIPPS CLIN., LA JOLLA, CALIF. 92037, USA.
MOL CELL BIOL 6 (12). 1986. 4329-4334. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/75 (Item 15 from file: 5)
5761770 BIOSIS Number: 83024077
MATING TYPE-LIKE CONVERSION PROMOTED BY THE 2 MICROMETER CIRCLE SITE-SPECIFIC RECOMBINASE IMPLICATIONS FOR THE DOUBLE-STRAND-GAP REPAIR MODEL
JAYARAM M
DEP. MOLECULAR BIOLOGY, RESEARCH INST. SCRIPPS CLINIC, LA JOLLA, CALIFORNIA 92037.
MOL CELL BIOL 6 (11). 1986. 3831-3837. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/76 (Item 16 from file: 5)
5751545 BIOSIS Number: 83013852
ASSOCIATION OF RECIPROCAL EXCHANGE WITH GENE CONVERSION BETWEEN THE REPEATED SEGMENTS OF 2-MICROMETER CIRCLE
JAYARAM M
DEPARTMENT OF MOLECULAR BIOLOGY, RESEARCH INSTITUTE OF SCRIPPS CLINIC, 10666 NORTH TORREY PINES ROAD, LA JOLLA, CALIF. 92037, USA.
J MOL BIOL 191 (3). 1986. 341-354. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/77 (Item 17 from file: 5)
5696494 BIOSIS Number: 33091515
MECHANISMS OF ACTION OF THE FLP RECOMBINASE OF THE 2-MICRON PLASMID OF YEAST
SADOWSKI P D; BEATTY L G; CLARY D; OLLERHEAD S
DEP. MED. GENETICS, MED. SCIENCES BUILD., UNIV. TORONTO, TORONTO, CANADA M5S 1A8.
MCMACKEN, R. AND T. J. KELLY (ED.). UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY NEW SERIES, VOL. 47. DNA REPLICATION AND RECOMBINATION; PARK CITY, UTAH, USA, MARCH 16-23, 1986. XXVI+782P. ALAN R. LISS, INC.: NEW YORK, NEW YORK, USA. ILLUS. ISBN 0-8451-2646-6. 0 (0). 1987. 691-702. CODEN: USMBD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/78 (Item 18 from file: 5)
5504855 BIOSIS Number: 32027162
INTERACTION OF THE FLP RECOMBINASE OF THE 2-MICRON PLASMID WITH ITS TARGET SEQUENCE
SADOWSKI P D; ANDREWS B J; BEATTY L G; SIDENBERG D; PROTEAU G
DEP. MED. GENETICS, UNIV. TORONTO, TORONTO M5S 1A8, CAN.
KLAR, A. AND J. N. STRATHERN (ED.). CURRENT COMMUNICATIONS IN MOLECULAR BIOLOGY: MECHANISMS OF YEAST RECOMBINATION; MEETING, COLD SPRING HARBOR, N.Y., USA. IX+193P. COLD SPRING HARBOR LABORATORY: COLD SPRING HARBOR, N.Y., USA. ILLUS. PAPER. ISBN 0-87969-195-6. 0 (0). 1986. 7-10. CODEN: 24607
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/79 (Item 19 from file: 5)
5426144 BIOSIS Number: 82070947
INTERACTION OF THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2 MICROMETER PLASMID WITH MUTATED TARGET SEQUENCES
ANDREWS B J; MCLEOD M; BROACH J; SADOWSKI P D
DEP. OF MED. GENETICS, UNIV. OF TORONTO, TORONTO, ONTARIO M5S 1A8, CANADA.
MOL CELL BIOL 6 (7). 1986. 2482-2489. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/80 (Item 20 from file: 5)
5389362 BIOSIS Number: 82034165
FLP SITE-SPECIFIC RECOMBINASE OF YEAST 2-MICROMETER PLASMID TOPOLOGICAL FEATURES OF THE REACTION
BEATTY L G; BABINEAU-CLARY D; HOGREFE C; SADOWSKI P D
DEP. OF MED. GENETICS, UNIV. OF TORONTO, TORONTO M5S 1A8, CANADA.
J MOL BIOL 188 (4). 1986. 529-544. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/81 (Item 21 from file: 5)
5265813 BIOSIS Number: 81033120
THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2 MICROMETER PLASMID ATTACHES COVALENTLY TO DNA VIA A PHOSPHOTYROSYL LINKAGE
GRONOSTAJSKI R M; SADOWSKI P D
DEP. MED. GENET., UNIV. TORONTO, TORONTO, ONT. M5S1A8, CAN.
MOL CELL BIOL 5 (11). 1985. 3274-3279. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/82 (Item 22 from file: 5)
5256098 BIOSIS Number: 81023405
THE FLP PROTEIN OF THE 2-MICRON PLASMID OF YEAST SACCHAROMYCES-CEREVISIAE PURIFICATION OF THE PROTEIN FROM ESCHERICHIA-COLI CELLS EXPRESSING THE CLONED FLP GENE
BABINEAU D; VETTER D; ANDREWS B J; GRONOSTAJSKI R M; PROTEAU G A; BEATTY L G; SADOWSKI P D
DEP. MED. GENETICS, UNIV. TORONTO, TORONTO, M5S 1A8, CANADA.

J BIOL CHEM 260 (22). 1985. 12313-12319. CODEN: JBCHA
Full Journal Title: Journal of Biological Chemistry
Language: ENGLISH

2/3/83 (Item 23 from file: 5)
5168213 BIOSIS Number: 31057528
THE FLP RECOMBINASE OF THE 2-MICRON PLASMID OF YEAST
SADOWSKI P D; ANDREWS B J; BABINEAU-CLARY D; BEATTY L; GRONOSTAJSKI R M; PROTEAU G; SIDENBERG D
DEP. MED. GENET., UNIV. TORONTO, TORONTO M5S 1A8, CANADA.
SYMPOSIUM ON MECHANISMS OF DNA REPLICATION AND RECOMBINATION HELD AT THE 15TH ANNUAL MEETING OF THE UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, MAR. 16-23, 1986. J CELL BIOCHEM SUPPL 0 (10 PART B). 1986. 137. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/84 (Item 24 from file: 5)
4696890 BIOSIS Number: 29054205
INTERACTION OF THE FLP RECOMBINASE WITH SUBSTRATE 2-MICRON CIRCLE DNA
ANDREWS B J; BEATTY L; SADOWSKI P D
UNIV. TORONTO.
SYMPOSIUM ON YEAST CELL BIOLOGY HELD AT THE 14TH ANNUAL MEETING OF THE UCLA (UNIVERSITY OF CALIFORNIA - LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, APR. 9-15, 1985. J CELL BIOCHEM SUPPL 0 (9 PART C). 1985. 117. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/85 (Item 1 from file: 399)
116167825 CA: 116(17)167825y PATENT
Methods for in vitro recombination of multigene families for generation of new phenotypes
INVENTOR(AUTHOR): Short, Jay M.; Sorge, Joseph A.
LOCATION: USA
ASSIGNEE: Stratagene
PATENT: PCT International ; WO 9116427 A1 DATE: 911031
APPLICATION: WO 91US2910 (910424) *US 513957 (900424)
PAGES: 204 pp. CODEN: PIXXD2 LANGUAGE: English CLASS: C12N-015/00A; C12P-019/34B; C12P-021/06B; C07H-021/00B DESIGNATED COUNTRIES: AU; CA; FI; JP; KR; NO DESIGNATED REGIONAL: AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LU ; NL; SE

2/3/86 (Item 2 from file: 399)
106208826 CA: 106(25)208826p JOURNAL
Rapid localization and characterization of random mutations within the 2.mu. circle site-specific recombinase: a general strategy for analysis of protein function
AUTHOR(S): Govind, Nadathur S.; Jayaram, Makkuni
LOCATION: Res. Inst. Scripps Clin., La Jolla, CA, 92037, USA
JOURNAL: Gene DATE: 1987 VOLUME: 51 NUMBER: 1 PAGES: 31-41 CODEN: GENED6 ISSN: 0378-1119 LANGUAGE: English

2/3/87 (Item 3 from file: 399)
104001445 CA: 104(1)1445b JOURNAL
The FLP recombinase of the yeast 2-.mu.m plasmid: characterization of its recombination site
AUTHOR(S): Senecoff, Julie F.; Bruckner, Robert C.; Cox, Michael M.
LOCATION: Coll. Agric. Life Sci., Univ. Wisconsin, Madison, WI, 53706, USA
JOURNAL: Proc. Natl. Acad. Sci. U. S. A. DATE: 1985 VOLUME: 82
NUMBER: 21 PAGES: 7270-4 CODEN: PNASA6 ISSN: 0027-8424 LANGUAGE: English

2/3/88 (Item 4 from file: 399)
102216080 CA: 102(25)216080y JOURNAL
The FLP recombinase of the 2.mu. circle DNA of yeast: interaction with its target sequences
AUTHOR(S): Andrews, Brenda J.; Proteau, Gerald A.; Beatty, Linda G.; Sadowski, Paul D.
LOCATION: Dep. Med. Genet., Univ. Toronto, Toronto, ON, Can., M5S 1A8
JOURNAL: Cell (Cambridge, Mass.) DATE: 1985 VOLUME: 40 NUMBER: 4
PAGES: 795-803 CODEN: CELLB5 ISSN: 0092-8674 LANGUAGE: English

2/3/89 (Item 1 from file: 434)
11609410 Genuine Article#: HX635 No. References: 12
Title: HALF-SITE STRAND TRANSFER BY STEP-ARREST MUTANTS OF YEAST SITE-SPECIFIC RECOMBINASE FLP
Author(s): SERRE MC; JAYARAM M
Corporate Source: UNIV TEXAS,DEPT MICROBIOL/AUSTIN//TX/78712; UNIV TEXAS,DEPT MICROBIOL/AUSTIN//TX/78712
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1992, V225, N3 (JUN 5), P643-649
Language: ENGLISH Document Type: ARTICLE

2/3/90 (Item 2 from file: 434)
11609409 Genuine Article#: HX635 No. References: 25
Title: HALF-SITE RECOMBINATIONS MEDIATED BY YEAST SITE-SPECIFIC RECOMBINASE-FLP AND RECOMBINASE-R
Author(s): SERRE MC; EVANS BR; ARAKI H; OSHIMA Y; JAYARAM M
Corporate Source: UNIV TEXAS,DEPT MICROBIOL/AUSTIN//TX/78712; UNIV TEXAS,DEPT MICROBIOL/AUSTIN//TX/78712; OSAKA UNIV,FAC ENGN,DEPT FERMENTAT TECHNOL/SUITA/OSAKA 565/JAPAN/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1992, V225, N3 (JUN 5), P621-642
Language: ENGLISH Document Type: ARTICLE

2/3/91 (Item 3 from file: 434)
11603498 Genuine Article#: HX080 No. References: 40
Title: MUTAGENESIS OF A CONSERVED REGION OF THE GENE ENCODING THE FLP RECOMBINASE OF SACCHAROMYCES-CEREVISIAE - A ROLE FOR ARGININE-191 IN BINDING AND LIGATION

Author(s): FRIESEN H; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MOLEC & MED GENET/TORONTO M5S1A8/ONTARIO/CANADA/; UNIV TORONTO,DEPT MOLEC & MED GENET/TORONTO M5S1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1992, V225, N2 (MAY 20), P313-326
Language: ENGLISH Document Type: ARTICLE

2/3/92 (Item 4 from file: 434)
11588831 Genuine Article#: HV855 No. References: 41
Title: SITE-SPECIFIC RECOMBINASE, R, ENCODED BY YEAST PLASMID PSR1
Author(s): ARAKI H; NAKANISHI N; EVANS BR; MATSUZAKI H; JAYARAM M; OSHIMA Y
Corporate Source: OSAKA UNIV,FAC ENGN,DEPT BIOTECHNOL,2-1 YAMADAOKA/SUITA/OSAKA 565/JAPAN/; OSAKA UNIV,FAC ENGN,DEPT BIOTECHNOL,2-1 YAMADAOKA/SUITA/OSAKA 565/JAPAN/; UNIV TEXAS,DEPT MICROBIOL/AUSTIN//TX/78712
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1992, V225, N1 (MAY 5), P25-37
Language: ENGLISH Document Type: ARTICLE

2/3/93 (Item 5 from file: 434)
11506141 Genuine Article#: HN234 No. References: 35
Title: SITE-SPECIFIC RECOMBINATION OF 2-MU-M PLASMID OF YEAST SACCHAROMYCES-CEREVISIAE
Author(s): PUSHNOVA EA
Corporate Source: ST PETERBURG PEDIAT MED INST/ST PETERBURG//USSR/
Journal: GENETIKA, 1992, V28, N2 (FEB), P25-34
Language: RUSSIAN Document Type: ARTICLE (Abstract Available)

2/3/94 (Item 6 from file: 434)
11487805 Genuine Article#: HM053 No. References: 33
Title: SITE-SPECIFIC INTEGRATION OF THE HAEMOPHILUS-INFLUENZAE BACTERIOPHAGE HP1 - IDENTIFICATION OF THE POINTS OF RECOMBINATIONAL STRAND EXCHANGE AND THE LIMITS OF THE HOST ATTACHMENT SITE
Author(s): HAUSER MA; SCOCCA JJ
Corporate Source: JOHNS HOPKINS UNIV,SCH HYG & PUBL HLTH,DEPT BIOCHEM/BALTIMORE//MD/21205; JOHNS HOPKINS UNIV,SCH HYG & PUBL HLTH,DEPT BIOCHEM/BALTIMORE//MD/21205
Journal: JOURNAL OF BIOLOGICAL CHEMISTRY, 1992, V267, N10 (APR 5), P 6859-6864
Language: ENGLISH Document Type: ARTICLE (Abstract Available)

2/3/95 (Item 7 from file: 434)
11338662 Genuine Article#: HB304 No. References: 21
Title: EXCHANGE OF GENE ACTIVITY IN TRANSGENIC PLANTS CATALYZED BY THE CRE-LOX SITE-SPECIFIC RECOMBINATION SYSTEM
Author(s): BAYLEY CC; MORGAN M; DALE EC; OW DW
Corporate Source: USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; UNIV CALIF BERKELEY,DEPT PLANT PATHOL/BERKELEY//CA/94720
Journal: PLANT MOLECULAR BIOLOGY, 1992, V18, N2 (JAN), P353-361
Language: ENGLISH Document Type: ARTICLE (Abstract Available

2/3/96 (Item 8 from file: 434)
11317754 Genuine Article#: GZ516 No. References: 33

Title: A FROG VIRUS-3 GENE CODES FOR A PROTEIN CONTAINING THE MOTIF CHARACTERISTIC OF THE INT FAMILY OF INTEGRASES
Author(s): ROHOZINSKI J; GOORHA R
Corporate Source: ST JUDE CHILDRENS HOSP,DEPT VIROL & MOLEC BIOL,332 N LAUDERDALE,POB 318/MEMPHIS//TN/38101; ST JUDE CHILDRENS HOSP,DEPT VIROL & MOLEC BIOL,332 N LAUDERDALE,POB 318/MEMPHIS//TN/38101
Journal: VIROLOGY, 1992, V186, N2 (FEB), P693-700
Language: ENGLISH Document Type: ARTICLE

2/3/97 (Item 9 from file: 434)
10583597 Genuine Article#: EP811 No. References: 61
Title: A NOVEL RECOMBINATOR IN YEAST BASED ON GENE-II PROTEIN FROM BACTERIOPHAGE-F1
Author(s): STRATHERN JN; WEINSTOCK KG; HIGGINS DR; MCGILL CB
Corporate Source: NCI,FREDERICK CANC RES & DEV CTR,BASIC RES PROGRAM/FREDERICK//MD/21701
Journal: GENETICS, 1991, V127, N1, P61-73
Language: ENGLISH Document Type: ARTICLE (Abstract Available)

2/3/98 (Item 10 from file: 434)
09323349 Genuine Article#: T4208 No. References: 45
Title: FLP RECOMBINASE OF THE 2-MU-M CIRCLE PLASMID OF SACCHAROMYCES-CEREVISIAE BENDS ITS DNA TARGET - ISOLATION OF FLP MUTANTS DEFECTIVE IN DNA BENDING
Author(s): SCHWARTZ CJE; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1989, V205, N4, P647-658
Language: ENGLISH Document Type: ARTICLE

2/3/99 (Item 11 from file: 434)
07863892 Genuine Article#: F8861 No. References: 37
Title: ISOLATION OF INTERMEDIATES IN THE BINDING OF THE FLP RECOMBINASE OF THE YEAST PLASMID 2-MIRON CIRCLE TO ITS TARGET SEQUENCE
Author(s): ANDREWS BJ; BEATTY LG; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1987, V193, N2, P345-358
Language: ENGLISH Document Type: ARTICLE

2/3/100 (Item 12 from file: 434)
07372665 Genuine Article#: C9356 No. References: 23
Title: INTERACTION OF THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2-MU-M PLASMID WITH MUTATED TARGET SEQUENCES
Author(s): ANDREWS BJ; MCLEOD M; BROACH J; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/; COLD SPRING HARBOR LAB/COLD SPRING HARBOR//NY/11724; PRINCETON UNIV,DEPT MOLEC BIOL/PRINCETON//NJ/08544
Journal: MOLECULAR AND CELLULAR BIOLOGY, 1986, V6, N7, P2482-2489
Language: ENGLISH Document Type: ARTICLE

2/3/101 (Item 13 from file: 434)
07260459 Genuine Article#: C1205 No. References: 44
Title: FLP SITE-SPECIFIC RECOMBINASE OF YEAST 2-MU-M PLASMID - TOPOLOGICAL

FEATURES OF THE REACTION
Author(s): BEATTY LG; BABINEAUCLARY D; HOGREFE C; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1986, V188, N4, P529-544
Language: ENGLISH Document Type: ARTICLE

2/3/102 (Item 14 from file: 434)
06806789 Genuine Article#: AUF29 No. References: 22
Title: THE FLP RECOMBINASE OF THE YEAST 2-MU-M PLASMID - CHARACTERIZATION OF ITS RECOMBINATION SITE
Author(s): SENECOFF JF; BRUCKNER RC; COX MM
Corporate Source: UNIV WISCONSIN,COLL AGR & LIFE SCI,DEPT BIOCHEM,420 HENRY MALL/MADISON//WI/53706
Journal: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 1985, V82, N21, P7270-7274
Language: ENGLISH Document Type: ARTICLE

2/3/103 (Item 15 from file: 434)
06780315 Genuine Article#: ATE60 No. References: 28
Title: THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2-MU-M PLASMID ATTACHES COVALENTLY TO DNA VIA A PHOSPHOTYROSYL LINKAGE
Author(s): GRONOSTAJSKI RM; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: MOLECULAR AND CELLULAR BIOLOGY, 1985, V5, N11, P3274-3279
Language: ENGLISH Document Type: ARTICLE

2/3/104 (Item 1 from file: 440)
03761331 Genuine Article#: HZ483 No. References: 12
Title: LIGATION ACTIVITY OF FLP RECOMBINASE - THE STRAND LIGATION ACTIVITY OF A SITE-SPECIFIC RECOMBINASE USING AN ACTIVATED DNA SUBSTRATE
Author(s): PAN GH; SADOWSKI PD (Reprint)
Corporate Source: UNIV TORONTO,DEPT MOLEC & MED GENET/TORONTO M5S1A8/ONTARIO/CANADA/ (Reprint); UNIV TORONTO,DEPT MOLEC & MED GENET/TORONTO M5S1A8/ONTARIO/CANADA/
Journal: JOURNAL OF BIOLOGICAL CHEMISTRY, 1992, V267, N18 (JUN 25), P 12397-12399
Language: ENGLISH Document Type: NOTE (Abstract Available)

2/3/105 (Item 1 from file: 76)
1171271 82001618771
Mutations in the 2-.mu.m circle site-specific recombinase that abolish recombination without affecting substrate recognition.
Prasad, P.V.; Young, L.-J.; Jayaram, M.
Dep. Mol. Biol., Res. Inst. Scripps Clin., 10666 N. Torrey Pines Rd., La Jolla, CA 92037, USA
PROC. NATL. ACAD. SCI. USA; 84(8), pp. 2189-2193 1987
Language: English Summary Language: English

2/3/106 (Item 1 from file: 73)
8210454 EMBASE No: 91239554
Erratum: Identification of the active site tyrosine of Flp recombinase. Possible relevance of its location to the mechanism of recombination (Vol.

265 (1990) 18504-18510)
Evans B.R.; Chen J.-W.; Parsons R.L.; Bauer T.K.; Teplow D.B.; Jayaram M.
J. BIOL. CHEM. (USA) , 1991, 266/11 (7312) CODEN: JBCHA ISSN: 0021-9258
LANGUAGES: English

2/3/107 (Item 2 from file: 73)
7363228 EMBASE No: 89079376
FLP recombinase of the 2 microm circle plasmid of Saccharomyces cerevisiae bends its DNA target. Isolation of FLP mutants defective in DNA bending
Schwartz C.J.E.; Sadowski P.D.
Department of Medical Genetics, University of Toronto, Toronto, Ont. M5S 1A8 Canada
J. MOL. BIOL. (United Kingdom) , 1989, 205/4 (647-658) CODEN: JMOBA ISSN: 0022-2836
LANGUAGES: English

2/3/108 (Item 1 from file: 144)
09775158 PASCAL No.: 91-0572331
Domain of a yeast site-specific recombinase (Flp) that recognizes its target site
JING-WEN CHEN; EVANS B R; SANG-HWA YANG; TEPLOW D/ B; JAYARAM M
Univ. Texas, dep. microbiology, Austin TX 78712, USA
Journal: Proceedings of the National Academy of Sciences of the United States of America, 1991, 88 (14) 5944-5948
Language: English

2/3/109 (Item 2 from file: 144)
09771721 PASCAL No.: 91-0568894
Protein-based assymmetry and protein-protein interactions in FLP recombinase-mediated site-specific recombination
XIAO-HONG QIAN; INMAN R B; COX M M
Univ. Wisconsin, coll. agricultural life sci., dep. biochemistry, Madison WI 53706, USA
Journal: Journal of biological chemistry (The), 1990, 265 (35) 21779-21788
Language: English

2/3/110 (Item 3 from file: 144)
09730857 PASCAL No.: 91-0527991
Site-specific recombination between homologous chromosomes in Drosophila
GOLIC K G
Univ. Chicago, Howard Hughes medical inst., dep; molcular genetics cell biology, Chicago IL 60637, USA
Journal: Science : (Washington, DC), 1991, 252 (5008) 958-961
Language: English

2/3/111 (Item 4 from file: 144)
09563896 PASCAL No.: 91-0354326
Tyr60 variants of Flp recombinase generate conformationally altered protein-DNA complexes : differential activity in full-site and half-site recombinations
JING-WEN CHEN; EVANS B R; LEI ZHENG; JAYARAM M

Univ. Texas at Austin, dep. microbiology, Austin TX 78712, USA
Journal: Journal of molecular biology, 1991, 218 (1) 107-118
Language: English

2/3/112 (Item 5 from file: 144)
07823248 PASCAL No.: 87-0302971
Interaction of the FLP recombinase of the saccharomyces cerevisiae 2 mu m plasmid with mutated target sequences
NDREWS B J; MCLEOD M; BROACH J; SADOWSKI P D
Univ. Toronto, dep. medical genetics, Toronto ON M5S 1A8, Canada
Journal: Molecular and cellular biology, 1986, 6 (7) 2482-2489
Language: ENGLISH

2/3/113 (Item 1 from file: 77)
89015048 V17N02
FLP recombinase induction of the breakage-fusion-bridge cycle (BFBC) and gene conversion in Saccharomyces cerevisiae
Rank, G.H.; Xiao, W.; Kolenovsky, A.; Arndt, G.
Univ. Saskatchewan, Saskatoon, Sask., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome Poster Paper

2/3/114 (Item 2 from file: 77)
89014585 V17N02
Structure-function relationship of the sequence specific DNA binding function of the FLP recombinase
Amin, A.A.; Sadowski, P.D.
Univ. Toronto, Toronto, Ont., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome Poster Paper

2/3/115 (Item 3 from file: 77)
89014584 V17N02
FLP recombinase of 2 mu circle of S. cerevisiae bends its DNA target: An in vitro analysis
Schwartz, C.J.E.; Sadowski, P.D.
Univ. Toronto, Toronto, Ont., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome Poster Paper

2/3/116 (Item 4 from file: 77)
89013277 V17N02

Mutational analysis of the FLP site-specific recombinase of the yeast 2 micron plasmid
Sadowski, P.
Univ. Toronto, Toronto, Ont., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome

2/3/117 (Item 5 from file: 77)
89012894 V17N02
Step-arrest mutants of FLP recombinase: Implications for the mechanism of recombination
Evans, B.R.; Parsons, R.; Crain, K.; Jayaram, M.
Mol. Biol. Dep., Res. Inst. Scripps Clin. and Res. Found., La Jolla, CA, USA
14th International Conference on Yeast Genetics and Molecular Biology 8830578 Espoo (Finland) 7-13 Aug 1988
European Association for Cancer Research
Subscription Department C, John Wiley & Sons Inc., 605 Third Avenue, New York, NY 10158 (USA), Abstracts will be Published in Special Issue of Journal 'Yeast' Volume 4. ISSN 0749-503X

2/3/118 (Item 1 from file: 265)
0128688 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 5R01GM35654-07 AGENCY CODE: CRISP
Site specific recombination in the yeast plasmid 2 micron circle
PRINCIPAL INVESTIGATOR: JAYARAM, MAKKUNI
ADDRESS: UNIVERSITY OF TEXAS DEPT OF MICROBIOLOGY AUSTIN, TX 78712
PERFORMING ORG.: UNIVERSITY OF TEXAS AUSTIN, AUSTIN, TEXAS
SPONSORING ORG.: NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
FY : 92 FUNDS: $265,024

2/3/119 (Item 2 from file: 265)
0092015 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 1R01HD28694-01 AGENCY CODE: CRISP
Site-specific recombination in spermatogenesis (Drosophila)
PRINCIPAL INVESTIGATOR: GOLIC, KENT G
ADDRESS: UNIVERSITY OF UTAH SALT LAKE CITY, UT 84112
PERFORMING ORG.: UNIVERSITY OF UTAH, SALT LAKE CITY, UTAH
SPONSORING ORG.: NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT
FY : 92 FUNDS: $152,007

2/3/120 (Item 3 from file: 265)
0019654 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 9105934; 9105934 AGENCY CODE: NSF
Genetic Analysis of Pattern Formation During Drosophila Neurogenesis
PRINCIPAL INVESTIGATOR: Ellis, Hilary M Dr.
PERFORMING ORG.: Emory University, Biology, Atlanta, GA 30322
PROJECT MONITOR: Data is not available
SPONSORING ORG.: National Science Foundation, DIV OF INTEGRATIVE BIOLOGY & NEUROSCIENC, Washington, D.C., 20550

DATES: 910715 TO 920630 FY : 91 FUNDS: $69,613

2/3/121 (Item 4 from file: 265)
0019101 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 9103946; 9103946 AGENCY CODE: NSF
Generation of Mosaicism in Mice by a Site-Specific Recombinase (FLP)
PRINCIPAL INVESTIGATOR: O'Gorman, Stephen Dr.
PERFORMING ORG.: Salk Institute for Biological Studies, Gene Expression Laboratory, San Diego, CA 92128
PROJECT MONITOR: Thomas E. Brady
SPONSORING ORG.: National Science Foundation, DIV OF INTEGRATIVE BIOLOGY & NEUROSCIENC, Washington, D.C., 20550
DATES: 910315 TO 920831 FY : 91 FUNDS: $49,522

2/3/122 (Item 5 from file: 265)
0016053 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 9019220; 9019220 AGENCY CODE: NSF
Genetic Analysis in Arabidopsis
PRINCIPAL INVESTIGATOR: Signer, Ethan R Dr.
PERFORMING ORG.: Massachusetts Institute of Technology, Biology, Cambridge, MA 02139
PROJECT MONITOR: DeLill Nasser
SPONSORING ORG.: National Science Foundation, DIV OF MOLECULAR & CELLULAR BIOSCIENCES, Washington, D.C., 20550
DATES: 910201 TO 930731 FY : 91 FUNDS: $200,000

2/3/123 (Item 1 from file: 35)
01212062 ORDER NO: AADNN-59965
THE ROLE OF DNA BENDING IN FLP-MEDIATED SITE-SPECIFIC RECOMBINATION
Author: SCHWARTZ, CAROL JUDITH ELAINE
Degree: PH.D.
Year: 1990
Corporate Source/Institution: UNIVERSITY OF TORONTO (CANADA) (0779)
Source: VOLUME 52/11-B OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 5647. 209 PAGES
ISBN: 0-315-59965-0

2/3/124 (Item 2 from file: 35)
01142876 ORDER NO: AAD90-30816
UNUSUAL DNA STRUCTURE IN SITE-SPECIFIC AND HOMOLOGOUS RECOMBINATION (RECOMBINATION)
Author: UMLAUF, SCOTT W.
Degree: PH.D.
Year: 1990
Corporate Source/Institution: THE UNIVERSITY OF WISCONSIN - MADISON (0262)
Source: VOLUME 51/09-B OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 4199. 219 PAGES

2/3/125 (Item 3 from file: 35)
1061565 ORDER NO: AAD89-12817
ANALYSIS OF THE MAJOR DNASE I HYPERSENSITIVE SITE ON THE YEAST TWO-MICRON DNA PLASMID
Author: STRAND, ANDREW DAVID

Degree: PH.D.
Year: 1989
Corporate Source/Institution: UNIVERSITY OF MINNESOTA (0130)
Source: VOLUME 50/02-B OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 446. 111 PAGES

2/3/126 (Item 4 from file: 35)
949308 ORDER NO: AAD87-06690
A GENETIC ANALYSIS OF FACTORS INVOLVED IN THE MAINTENANCE OF THE 2 MICRON PLASMID OF SACCHAROMYCES CEREVISIAE (CHROMATIN)
Author: VEIT, BRUCE EDWARD
Degree: PH.D.
Year: 1986
Corporate Source/Institution: UNIVERSITY OF WASHINGTON (0250)
Source: VOLUME 47/12-B OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 4763. 97 PAGES

2/3/127 (Item 1 from file: 51)
00405585 91-03-b0028 SUBFILE: FSTA
Yeast 2 MUm vectors replicate and undergo recombination in Torulaspora delbrueckii.
Compagno, C.; Ranzi, B. M.; Martegani, E.
Correspondence (Reprint) address, B. M. Ranzi, Dipartimento di Fisiologia e Biochimica Generali, Sezione di Biochimica Comparata, Univ. di Milano, Milan, Italy
Molecular Microbiology 1989 , 3 (8) 1003-1010
LANGUAGE: English

2/3/128 (Item 1 from file: 60)
09154644
PROJ NO: NYC-186301 AGENCY : SAES NY.C
PROJ TYPE: STATE
START: 01 JUL 91 TERM: 30 JUN 92
INVEST: MACINTYRE R J
ENTOMOLOGY
CORNELL UNIVERSITY
ITHACA NEW YORK 14853

DEVELOPMENT OF A MORE EFFICIENT INSECT TRANSFORMATION SYSTEM

OBJECTIVES:The goal of the research described below is to develop a system in which DNA canbe both easily and effectively delivered to insect embryos and, using the yeast "flip recombinase" system, insure the recovery of transgenic animals at high frequencies.

PRIMARY HEADINGS: R207 Insect Control-Field Crops;A4500 Protection Against Insects;C6500 Invertebrates;F1313 Physiology-Other

2/3/129 (Item 2 from file: 60)
09091400
PROJ NO: WIS02827 AGENCY : SAES WIS
PROJ TYPE: STATE
START: 01 JUL 86 TERM: 30 NOV 96 FY: 1989
INVEST: COX M M

BIOCHEMISTRY
UNIV OF WISCONSIN
MADISON WISCONSIN 53706

THE BIOCHEMISTRY OF GENETIC RECOMBINATION

OBJECTIVES:The FLP recombinase (derived from yeast) has been purified extensively. The properties of this protein and the recombination event it catalyzes are being studied in vitro. The recombination site utilized by this protein has been defined in detail. Studies on the mechanism of action of this recombination system are now getting underway.

PRIMARY HEADINGS: R318 Noncommodity Biotechnology, Biometry; A7000 Experimental Design, Statistical Methods;C6300 Biological Cell Systems; F0114 Biochemistry and Biophysics-Other

2/3/130 (Item 1 from file: 286)
0050984 Journal Announcement: 08APR91 Doc Type: 2
Nature, 15 MAR 1991, Vol(No) 251(4999), Page(s) 1351-1355

1ST COMPANY/ORGANIZATION NAME:
Salk Institute for Biological Studies, The, USA (1921)
?

?

?

?ds

Set Items Description
S1 365 FLP(10N)RECOMBINAS?
S2 130 RD (unique items)
?s s2 and mammal?

Processing
Processing
Processing
Processing
Processing
Processing
Processing
Processed 10 of 25 files ...
Processing
Processed 20 of 25 files ...
130 S2
2416144 MAMMAL?
S3 3 S2 AND MAMMAL?
d
?
?t3/3/1-3

3/3/1 (Item 1 from file: 155)
07645850 91164850
Recombinase-mediated gene activation and site-specific integration in mammalian cells.
O'Gorman S; Fox DT; Wahl GM
Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037.
Science Mar 15 1991, 251 (4999) p1351-5, ISSN 0036-8075
Journal Code: UJ7
Languages: ENGLISH
Document type: JOURNAL ARTICLE

3/3/2 (Item 1 from file: 434)
11338662 Genuine Article#: HB304 No. References: 21
Title: EXCHANGE OF GENE ACTIVITY IN TRANSGENIC PLANTS CATALYZED BY THE CRE-LOX SITE-SPECIFIC RECOMBINATION SYSTEM
Author(s): BAYLEY CC; MORGAN M; DALE EC; OW DW
Corporate Source: USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; UNIV CALIF BERKELEY,DEPT PLANT PATHOL/BERKELEY//CA/94720
Journal: PLANT MOLECULAR BIOLOGY, 1992, V18, N2 (JAN), P353-361
Language: ENGLISH Document Type: ARTICLE (Abstract Available)

3/3/3 (Item 1 from file: 286)
0050984 Journal Announcement: 08APR91 Doc Type: 2
Nature, 15 MAR 1991, Vol(No) 251(4999), Page(s) 1351-1355

1ST COMPANY/ORGANIZATION NAME:
Salk Institute for Biological Studies, Th , USA (1921)
?t3/4/1-3

3/4/1 (Item 1 from file: 155)
FN- DIALOG MEDLINE file 155
AN- 07645850|
AN- (NLM) 91164850|
TI- Recombinase-mediated gene activation and site-specific integration in mammalian cells.|
AU- O'Gorman S; Fox DT; Wahl GM|
CS- Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037.|
JN- Science; 251 (4999) p1351-5|
PY- Mar 15 1991|
SN- 0036-8075|
JC- UJ7|
LA- ENGLISH|
DT- JOURNAL ARTICLE|
JA- 9106|
SF- INDEX MEDICUS|
AB- A binary system for gene activation and site-specific integration, based on the conditional recombination of transfected sequences mediated by the FLP recombinase from yeast, was implemented in mammalian cells. In several cell lines, FLP rapidly and precisely

recombined copies of its specific target sequence to activate an otherwise silent beta-galactosidase reporter gene. Clones of marked cells were generated by excisional recombination within a chromosomally integrated copy of the silent reporter. By the reverse reaction, integration of transfected DNA was targeted to a specific chromosomal site. The results suggest that FLP could be used to mosaically activate or inactivate transgenes for analysis of vertebrate development, and to efficiently integrate transfected DNA at predetermined chromosomal locations.|
GS- Animal; In Vitro; Support, Non-U.S. Gov't|
DE- *DNA Nucleotidyltransferases/--Metabolism/--ME; /*Mammals/--Genetics/--GE; /*Recombination, Genetic; /*Transfection; beta-Galactosidase/--Genetics/--GE; /Animals, Transgenic; /Cell Line; /DNA Nucleotidyltransferases/--Genetics/--GE; /Restriction Mapping|
ID- EC 2.7.7.- (DNA Nucleotidyltransferases); EC 2.7.7.- (FLP recombinase); EC 3.2.1.23 (beta-Galactosidase)|
ID- FLP|

3/4/2 (Item 1 from file: 434)
FN- SCISEARCH_1974 - 9206W3
AN- 11338662|
GA- HB304|
TI- EXCHANGE OF GENE ACTIVITY IN TRANSGENIC PLANTS CATALYZED BY THE CRE-LOX SITE-SPECIFIC RECOMBINATION SYSTEM|
LA- ENGLISH|
AU- BAYLEY CC; MORGAN M; DALE EC; OW DW|
CS- USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; UNIV CALIF BERKELEY,DEPT PLANT PATHOL/BERKELEY//CA/94720|
GL- USA|
JN- PLANT MOLECULAR BIOLOGY, 1992, V18, N2, P353-361|
PY- 1992|
DT- ARTICLE|
NR- 21|
SF- SciSearch; CC LIFE--Current Contents, Life Sciences; CC AGRI--Current Contents, Agriculture, Biology & Environmental Sciences|
SC- BOTANY; BIOCHEMISTRY & MOLECULAR BIOLOGY|
AB- The Cre-lox site-specific recombination system of bacteriophage P1 was used to excise a firefly luciferase (luc) gene which had previously been incorporated into the tobacco genome. The excision event was due to site-specific DNA recombination between two lox sequences flanking the luc gene and was catalyzed by the Cre recombinase introduced by cross-fertilization. Recombination resulted in the fusion of a promoter with a distally located hygromycin phosphotransferase (hpt) coding sequence and the excision event was monitored as a phenotypic change from expression of luc to expression of hpt. The efficiency of recombination was estimated from the exchange of gene activity and confirmed by molecular analysis. The relevance to potential applications of site-specific deletion-fusion events for chromosome engineering are discussed.|
DE- Author Keywords: GENETIC ENGINEERING; PHAGE P1; RECOMBINASE; LUCIFERASE; SELECTABLE MARKERS|
ID- KeyWords Plus: FIREFLY LUCIFERASE GENE; FLP RECOMBINASE; MAMMALIAN-CELLS; 2-MU CIRCLE; DNA; YEAST; BACTERIOPHAGE-P1;

GENOME; EXPRESSION; SEQUENCES|
RF- 90-0047 002 (TRANSGENIC PLANTS; TRANSIENT EXPRESSION OF THE GUS GENE; INDICA RICE PROTOPLASTS; MICROPROJECTILE BOMBARDMENT; AGROBACTERIUM MEDIATED TRANSFORMATION)
90-1257 001 (BACILLUS-THURINGIENSIS STRAINS; TRANSGENIC TOBACCO PLANTS; DRY BEANS (PHASEOLUS-VULGARIS L); EXPRESSION OF INSECTICIDAL ACTIVITY; INSECT MIDGUT)
90-2362 001 (STA58 MAJOR ANTIGEN GENE; RHODOCOCCUS-FASCIANS CLONING VECTORS; ESCHERICHIA-COLI CHROMOSOME; PRECISE IDENTIFICATION)
90-4791 001 (FIREFLY LUCIFERASE EXPRESSION IN TRANSGENIC PLANTS; PROTEIN OF MAIZE TRANSPOSABLE ELEMENT AC; CAULIFLOWER MOSAIC-VIRUS; REPORTER GENES)
90-7783 001 (POLYMERASE CHAIN-REACTION; DNA AMPLIFICATION; POLYMORPHIC NUCLEOTIDE SUBSTITUTIONS IN BETA-GLOBIN GENES)
CR- ANDREWS BJ, 1985, V40, P795, CELL
BEVAN M, 1984, V12, P8711, NUCLEIC ACIDS RES
BROACH JR, 1980, V21, P501, CELL
CRAIG NL, 1988, V22, P77, ANNU REV GENET
DALE EC, 1990, V91, P79, GENE
DEWET JR, 1987, V7, P725, MOL CELL BIOL
GOLIC KG, 1989, V59, P499, CELL
HOESS RH, 1985, V181, P351, J MOL BIOL
HORSCH RB, 1985, V227, P1229, SCIENCE
KASTER KR, 1983, V11, P6895, NUCLEIC ACIDS RES
MANIATIS T, 1982, MOL CLONING
ODELL J, 1990, V223, P369, MOL GEN GENET
OGORMAN S, 1991, V251, P1351, SCIENCE
OW DW, 1986, V234, P856, SCIENCE
SAIKI RK, 1988, V239, P487, SCIENCE
SAUER B, 1989, V17, P147, NUCLEIC ACIDS RES
SAUER B, 1988, V85, P5166, P NATL ACAD SCI USA
SENECOFF JF, 1985, V82, P7270, P NATL ACAD SCI USA
STERNBERG N, 1981, V150, P467, J MOL BIOL
VAECK M, 1987, V328, P33, NATURE
VERWOERD TC, 1989, V17, P2362, NUCLEIC ACIDS RES||

3/4/3 (Item 1 from file: 286)
FN- DIALOG File 286: 286
AN- 00509841
JA- 08APR91|
DT- 2|
JN- Nature, 15 MAR 1991, Vol(No) 251(4999), Page(s) 1351-1355|
AB- Salk Institute scientists have shown that the site specific recombinase enzyme, FLP, from Saccharomyces cerevisiae can be used for gene activation in mammalian cells and have suggested it may be useful to mosaically inactivate or activate transgenes or to efficiently integrate transfected DNA at predetermined chromosomal locations.|
CS- Salk Institute for Biological Studies, The, USA (1921)|
?logoff

=> d his

(FILE 'USPAT' ENTERED AT 09:57:49 ON 12 JUL 94)
SET PAGELENGTH SCROLL
L1 5 S (YEAST? OR CEREVISIAE?)(30A)(FRT OR FLP OR RECOMBINASE?)

=> d 1-5

1. 5,268,296, Dec. 7, 1993, DNA vector and recombinant host cell for production of hirullin P6 and P18; Reinhard Maschler, et al., 435/252.3, 69.1, 172.3, 320.1, 942; 536/23.5 [IMAGE AVAILABLE]

2. 5,268,285, Dec. 7, 1993, Strains of yeast with increased rates of glycolysis; David T. Rogers, et al., 435/172.3, 161, 194, 254.21, 320.1 [IMAGE AVAILABLE]

3. 5,227,288, Jul. 13, 1993, DNA sequencing vector with reversible insert; Frederick R. Blattner, 435/6, 252.3, 252.33, 320.1; 935/29, 72, 73 [IMAGE AVAILABLE]

4. 5,114,922, May 19, 1992, Polypeptides with an anticoagulant activity; Reinhard Maschler, et al., 514/12; 530/324 [IMAGE AVAILABLE]

5. 4,997,757, Mar. 5, 1991, Process for detecting potential carcinogens; Robert H. Schiestl, 435/172.1, 6, 29, 172.3; 935/76, 78, 79, 84 [IMAGE AVAILABLE]

=>

=> d his l2

(FILE 'USPAT' ENTERED AT 09:57:49 ON 12 JUL 94)
L2 1 S MAMMAL?(100A)(FRT OR FLP OR RECOMBINASE?)

=> d kwic

US PAT NO: 5,159,066 [IMAGE AVAILABLE] L2: 1 of 1

ABSTRACT:
Recombination activating gene of mammalian origin (RAG-1), cDNA of RAG-1 of mammalian origin, mRNA expressed by RAG-1, the encoded recombinase and antibodies specific for the recombinase, as well as the use of the same for a diagnostic or therapeutic purpose.

DETDESC:

DETD(2)

The present invention relates to a gene of mammalian origin, referred to as recombination activating gene or RAG-1, which confers the ability to carry out V(D)J recombination on cells in which it is expressed. The RAG-1 gene product is thus a direct or indirect activator of V(D)J recombinase activity. The invention also refers to RAG-1 mRNA and to the RAG-1 encoded product. RAG-1 has been shown in pre-B. . . as in all transfectants into which it has been introduced. This pattern of expression is that expected for the V(D)J recombinase and, therefore, RAG-1 appears to be a master controller of the development of the effector cells of the immune system.. .

CLAIMS:

CLMS(3)

3. Isolated DNA of mammalian origin encoding recombinase.

=>

```
=> d his

   (FILE 'USPAT' ENTERED AT 14:51:48 ON 03 FEB 94)
         SET PAGELENGTH SCROLL
L1       0 S MAMMAL?(20A)(FLP OR FRT)(20A)(TRANSFECT? OR TRANSFORM? OR R
L2       5 S YEAST(20A)(FRT OR FLP)
L3       0 S MAMMAL?(50A)(FLP OR FRT)
L4       0 S MAMMAL?(100A)(FRT OR FLP)
L5       0 S MAMMAL(50A)(YEAST OR CEREVISIAE)(50A)RECOMBINASE?
L6       0 S MAMMAL?(200A)(YEAST? OR FUNG? OR CEREVISIAE)(200A)RECOMBINA

=> file jpoabs
FILE 'JPOABS' ENTERED AT 14:58:17 ON 03 FEB 94

 **    ****************************
 *  JAPANESE PATENT ABSTRACTS    *
 *                               *
 * CURRENTLY, DATA IS LOADED THROUGH THE ABSTRACT PUBLICATION  *
 * DATE OF JULY 5, 1993                *
 * THE LATEST GROUPS RECEIVED ARE: C1078 E1392, M1438 & P1567. *
 ********************************


=> s l1
       769 MAMMAL?
        21 FLP
        12 FRT
        32 TRANSFECT?
     39194 TRANSFORM?
      2791 RECOMB?
L7       0 MAMMAL?(20A)(FLP OR FRT)(20A)(TRANSFECT? OR TRANSFORM? OR RECO
         MB?)

=>  s l2
      2510 YEAST
        12 FRT
        21 FLP
L8       0 YEAST(20A)(FRT OR FLP)

=> s l3
       769 MAMMAL?
        21 FLP
        12 FRT
L9       0 MAMMAL?(50A)(FLP OR FRT)

=> s l4
       769 MAMMAL?
        12 FRT
        21 FLP
L10      0 MAMMAL?(100A)(FRT OR FLP)
```

```
=> s l5
        355 MAMMAL
       2510 YEAST
        278 CEREVISIAE
          0 RECOMBINASE?
L11       0 MAMMAL(50A)(YEAST OR CEREVISIAE)(50A)RECOMBINASE?

=> log y
U.S. Patent & Trademark Office LOGOFF AT 14:58:54 ON 03 FEB 94
```

mammal ? (100 n)(FLP or FRT)

DIALOG one search update

?
t2/7/1-16

2/7/1 (Item 1 from file: 434)
12729441 Genuine Article#: MK280 Number of References: 51
Title: SITE-SPECIFIC RECOMBINASES - TOOLS FOR GENOME ENGINEERING
Author(s): KILBY NJ; SNAITH MR; MURRAY JAH
Corporate Source: UNIV CAMBRIDGE,INST BIOTECHNOL,TENNIS COURT RD/CAMBRIDGE
CB2 1QT//ENGLAND/
Journal: TRENDS IN GENETICS, 1993, V9, N12 (DEC), P413-421
ISSN: 0168-9525
Language: ENGLISH Document Type: REVIEW
Abstract: Site-specific recombinases, from bacteriophage and yeasts have been developed as novel tools for manipulating DNA both in the test-tube and in living organisms. We discuss the characteristics of these enzyme systems, review their application in genetic and developmental studies and speculate on their future potential for large-scale directed modifications of eukaryotic genomes.

2/7/2 (Item 2 from file: 434)
12101712 Genuine Article#: KM161 Number of References: 24
Title: LIGATION OF SYNTHETIC ACTIVATED DNA SUBSTRATES BY SITE-SPECIFIC RECOMBINASES AND TOPOISOMERASE-I
Author(s): PAN GH; LUETKE K; JUBY CD; BROUSSEAU R; SADOWSKI P
Corporate Source: UNIV TORONTO,DEPT MOLEC & MED GENET/TORONTO M5S1A8/ONTARIO/CANADA/; UNIV TORONTO,DEPT MOLEC & MED GENET/TORONTO
M5S1A8/ONTARIO/CANADA/; NATL RES COUNCIL CANADA,BIOTECHNOL RES INST,GENET ENGN SECT/MONTREAL H4P 2R2/QUEBEC/CANADA/
Journal: JOURNAL OF BIOLOGICAL CHEMISTRY, 1993, V268, N5 (FEB 15), P 3683-3689
ISSN: 0021-9258
Language: ENGLISH Document Type: ARTICLE
Abstract: The FLP protein of the 2-mum plasmid of Saccharomyces cerevisiae is a conservative site-specific recombinase that is involved in the amplification of the plasmid. This recombination reaction proceeds via the covalent attachment of the protein to the 3'-phosphoryl group at the site of the breaks through a phosphotyrosine linkage. We have recently developed an assay that measures FLP-mediated strand ligation independent of FLP-mediated cleavage and covalent attachment to the DNA. The substrate for ligation was produced by FLP-induced cleavage of the FLP recognition site followed by digestion with Pronase and was shown to contain (at least) a tyrosine residue at the 3'-PO4 terminus adjacent to the FLP cleavage sites.

We have now synthesized artificial substrates that bear a tyrosine residue on the 3'-PO4 of an appropriate oligonucleotide and find that this substrate is ligated as efficiently as the previous ligation substrates that were isolated after FLP cleavage of the substrate. Analogous substrates for other members of the integrase family of recombinases (lambda integrase protein, P1-Cre protein) as well as for mammalian topoisomerase I are also active as ligation substrates with their cognate protein. This class of activated substrates should be

useful in the study of breakage and reunion reactions involving DNA.

2/7/3 (Item 3 from file: 434)
11338662 Genuine Article#: HB304 Number of References: 21
Title: EXCHANGE OF GENE ACTIVITY IN TRANSGENIC PLANTS CATALYZED BY THE
CRE-LOX SITE-SPECIFIC RECOMBINATION SYSTEM
Author(s): BAYLEY CC; MORGAN M; DALE EC; OW DW
Corporate Source: USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; UNIV CALIF BERKELEY,DEPT PLANT PATHOL/BERKELEY//CA/94720
Journal: PLANT MOLECULAR BIOLOGY, 1992, V18, N2 (JAN), P353-361
Language: ENGLISH Document Type: ARTICLE
Abstract: The Cre-lox site-specific recombination system of bacteriophage P1 was used to excise a firefly luciferase (luc) gene which had previously been incorporated into the tobacco genome. The excision event was due to site-specific DNA recombination between two lox sequences flanking the luc gene and was catalyzed by the Cre recombinase introduced by cross-fertilization. Recombination resulted in the fusion of a promoter with a distally located hygromycin phosphotransferase (hpt) coding sequence and the excision event was monitored as a phenotypic change from expression of luc to expression of hpt. The efficiency of recombination was estimated from the exchange of gene activity and confirmed by molecular analysis. The relevance to potential applications of site-specific deletion-fusion events for chromosome engineering are discussed.

2/7/4 (Item 4 from file: 434)
10707719 Genuine Article#: FB733 Number of References: 25
Title: RECOMBINASE-MEDIATED GENE ACTIVATION AND SITE-SPECIFIC INTEGRATION
IN MAMMALIAN-CELLS
Author(s): OGORMAN S; FOX DT; WAHL GM
Corporate Source: SALK INST BIOL STUDIES,GENE EXPRESS LAB/LA JOLLA//CA/92037
Journal: SCIENCE, 1991, V251, N4999, P1351-1355
Language: ENGLISH Document Type: ARTICLE
Abstract: A binary system for gene activation and site-specific integration, based on the conditional recombination of transfected sequences mediated by the FLP recombinase from yeast, was implemented in mammalian cells. In several cell lines, FLP rapidly and precisely recombined copies of its specific target sequence to activate an otherwise silent beta-galactosidase reporter gene. Clones of marked cells were generated by excisional recombination within a chromosomally integrated copy of the silent reporter. By the reverse reaction, integration of transfected DNA was targeted to a specific chromosomal site. The results suggest that FLP could be used to mosaically activate or inactivate transgenes for analysis of vertebrate development, and to efficiently integrate transfected DNA at predetermined chromosomal locations.

2/7/5 (Item 5 from file: 434)
10161463 Genuine Article#: DE904 Number of References: 75

Title: A SITE-SPECIFIC SELF-CLEAVAGE REACTION PERFORMED BY A NOVEL RNA IN
NEUROSPORA MITOCHONDRIA
Author(s): SAVILLE BJ; COLLINS RA
Corporate Source: UNIV TORONTO,DEPT BOT/TORONTO M5S 3B2/ONTARIO/CANADA/;
UNIV TORONTO,CTR PLANT BIOTECHNOL/TORONTO M5S 3B2/ONTARIO/CANADA/
Journal: CELL, 1990, V61, N4, P685-696
Language: ENGLISH Document Type: ARTICLE

2/7/6 (Item 1 from file: 357)
142790 DBA Accession No.: 93-00842
Gene transfer - gene transmission by retro virus vector, yeast artificial chromosome, mouse zygote homologous recombination, Cre-recombinase method and Flp system (conference abstract)
AUTHOR: Wagner E F
CORPORATE SOURCE: Research Institute of Molecular Pathology (IMP), Dr Bohr-Gasse 7, A-1030, Vienna, Austria.
JOURNAL: Science (258, Suppl., 31-32) 1992 CODEN: SCIEAS
LANGUAGE: English
ABSTRACT: Applications and limitations of gene transfer techniques were discussed. The high efficiency of retro virus infection allows the introduction of genes into cells, e.g. hematopoietic cells. These viral systems provide a method for the generation of animal models for human blood diseases and for possible gene therapy applications. However, the use of yeast artificial chromosomes introduced into cells via DNA-lipid micelles, or the generation of large transgenes through homologous recombination in mouse zygotes, provide a much superior gene transfer system to viral vector systems. Gene transfer techniques are also being used to inactivate a given gene locus by gene targeting. Two new loss-of-function approaches have recently been developed: (1) using the Cre- recombinase; and (2) using the Flp system. These 2 new methods may allow tissue-specific and developmentally regulated gene inactivation in transgenic mice as a function of the site-specific recombinase action. (7 ref)

2/7/7 (Item 2 from file: 357)
141482 DBA Accession No.: 92-13974 PATENT
FLP-mediated gene modification in mammalian cell - vector with FLP-recombinase gene and recombination site for e.g. gene targeting, gene therapy or transgenic animal development research
PATENT ASSIGNEE: Salk-Inst.Biol.Stud. 1992
PATENT NUMBER: WO 9215694 PATENT DATE: 920917 WPI ACCESSION NO.: 92-331739 (9240)
PRIORITY APPLIC. NO.: US 666252 APPLIC. DATE: 910308
NATIONAL APPLIC. NO.: WO 92US1899 APPLIC. DATE: 920306
LANGUAGE: English
ABSTRACT: A new mammalian recombination system comprises Saccharomyces cerevisiae FLP-recombinase (or a gene encoding it) and DNA containing at least 1 FLP recombination target site. The following are also new: DNA containing at least 1 FLP recombination site, at least 1 restriction site, at least 1 selectable marker, a bacterial (and optionally a mammal or virus) replication origin; the new DNA inserted

into the FLP recombination target site, and with a 2nd FLP target site in tandem with the 1st; methods for assembly of functional genes for activation of expression in mammal cells, disrupting gene expression in a mammal cells, recovery of transfected DNA from the genome of a transfected organism, and precisely targeted integration of DNA into the genome of a host, all using FLP-recombinase; a transgenic non-human mammal containing at least 1 FLP recombination site in its genome; a method for analysis of mammal development, using the above transgenic mammal and a vector encoding FLP under the control of a conditional promoter, and a reporter gene. The system may also be useful in gene therapy. (49pp)

2/7/8 (Item 3 from file: 357)
089171 DBA Accession No.: 89-07162
Production and isolation of large quantities of monoclonal antibody using serum-free medium and fast protein liquid chromatography - hybridoma cell culture
AUTHOR: Stocks S J; +Brooks D E
CORPORATE SOURCE: Department of Pathology, 2211 Wesbrook Mall, Vancouver, V6T 1W5, Canada.
JOURNAL: Hybridoma (8, 2, 241-47) 1989 CODEN: HYBRDY
LANGUAGE: English
ABSTRACT: A method for the production and purification of monoclonal antibody (MAb) on a large scale is described. 2 Hybridoma lines were used to generate monoclonal antibodies in serum-free medium; a rat-rat hybridoma specific for a surface antigen of a hybrid mouse cell line, and a mouse-mouse hybridoma line specific for rat IgG2a. The serum free medium (RPM1-1640) was supplemented with 5 ug/ml cattle insulin, and incubation was at 37 deg. Both hybridoma lines became confluent within 10 days at maximum cell density. Large-quantities of MAbs were produced in the medium, and purification was easily accomplished within a working day at 4 deg to retain high MAb activity. Ammonium sulfate precipitation, which can cause activity loss, was avoided. The serum free medium was purified by ultrafiltration through an Amicon XM100A filter and fast protein liquid chromatography on a mono Q column with an ionic strength and pH elution gradient. Yields obtained were between 10-30 mg pure MAb/l. (13 ref)

2/7/9 (Item 1 from file: 149)
10844741 Dialog File 149: Health Periodicals Database
Use Format 9 for FULL TEXT
TITLE: Site-specific recombination between homologous chromosomes in Drosophila.
AUTHOR: Golic, Kent G.
JOURNAL: Science VOL.: v252 ISSUE: n5008 PAGINATION: p958(4)
PUBLICATION DATE: May 17, 1991
AVAILABILITY: FULL TEXT Online LINE COUNT: 00224
SOURCE FILE: MI File 47

2/7/10 (Item 1 from file: 399)
118074746 CA: 118(9)74746z PATENT
Site-specific integration and excision of transforming DNA in animal cells using the Flp recombinase of yeast
INVENTOR(AUTHOR): Wahl, Geoffrey M.; O'Gorman, Stephen V.

LOCATION: USA
ASSIGNEE: Salk Institute for Biological Studies
PATENT: PCT International ; WO 9215694 A1 DATE: 920917
APPLICATION: WO 92US1899 (920306) *US 666252 (910308)
PAGES: 54 pp. CODEN: PIXXD2 LANGUAGE: English CLASS: C12N-015/85A; C12N-005/16B; C07H-015/12B DESIGNATED COUNTRIES: CA; JP
DESIGNATED REGIONAL: AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LU; MC; NL; SE
SECTION:
CA203001 Biochemical Genetics
IDENTIFIERS: recombination transforming DNA flp recombinase
DESCRIPTORS:
Development,mammalian...
anal. of, developmental regulation of reporter genes in, recombination of transforming DNA using flp recombinase in relation to
Genetic element,promoter...
developmentally regulated, expression of reporter genes from, recombination of transforming DNA using flp recombinase in relation to
Animal cell line,CV-1... Animal cell line,F9... Animal cell line,293...
expression in, of gene for flp recombinase, site-specific recombination of transforming DNA in relation to
Enzymes,DNA-recombining...
flp, gene for, expression in animal cells of, for site-specific integration or excision of transforming DNA
Saccharomyces cerevisiae... Saccharomyces...
flp recombinase of, gene for, expression in animal cells of, for site-specific integration or excision of transforming DNA
Gene,microbial...
for flp recombinase, expression in animal cell culture of, for site-specific integration or excision of transforming DNA
Genetic element...
frt (flp recombinase target site), site-specific recombination of transforming DNA in animal cells via, expression of flp recombinase gene in relation to
Animal cell...
mammalian, site-specific recombination of transforming DNA in, flp recombinase and frt sites in
Deoxyribonucleic acid sequences...
of flp recombinase gene of Saccharomyces cerevisiae
Protein sequences...
of flp recombinase of Saccharomyces cerevisiae
Recombination,genetic, site-specific... Recombination,genetic, site-specific excisive...
of transforming DNA in animal cell culture, flp recombinase and frt sites in
Mammal...
transgenic, site-specific recombination of transforming DNA in, flp recombinase and frt sites in
CAS REGISTRY NUMBERS:
145752-47-2 amino acid sequence of, complete, and expression in animal cell culture of gene for
145752-45-0 nucleotide sequence of
145752-46-1 nucleotide sequence of, complete, and expression in animal cell culture of

145752-44-9 nucleotide sequence of, in transforming DNA for site-specific recombination of transforming DNA

2/7/11 (Item 1 from file: 265)
0103231 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 1R01HD30255-01 AGENCY CODE: CRISP
FATE MAPS OF EMBRYONIC GENE EXPRESSION IN MICE
PRINCIPAL INVESTIGATOR: O'GORMAN, STEPHEN V
ADDRESS: SALK INSTITUTE PO BOX 85800 SAN DIEGO, CA 92186-5800
PERFORMING ORG.: SALK INSTITUTE FOR BIOLOGICAL STUDIES, SAN DIEGO,
CALIFORNIA
SPONSORING ORG.: NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT
FY : 93 FUNDS: $260,269 TYPE OF AWARD: New Award (Type 1)
SUMMARY: The long term objective of the research initiated by this proposal is to investigate the genetic regulation of mammalian development. The principal experimental objective is to compile fate maps of the mature fates of cells that descend from progenitors that transiently express specific candidate mammalian developmental control genes during embryonic and fetal stages in transgenic mice. This will be done both in normal mice, and in mice that fail to express the normal products of genes of interest. A novel molecular paradigm for fate mapping will be employed that is based on the precise recombination of transgenes by the yeast recombinase FLP. By this means, the transient activity of a gene can be used to indelibly mark not only the cells in which the gene is expressed, but all of its descendants, even if the latter do not express the gene.

The specific aims of this proposal are to define the descendant domains (lineages) established by progenitors that transiently express Hox 2.9, Krox 20, or Hox 2.6 in the hindbrain and adjacent branchial arch tissues in both normal animals and in, animals that fail to express the normal products of these genes. Both the descendant expression domains and the descendant functional domains of these genes will be mapped and distinguished from one another. The first product of the research program will be a fate map of the mouse that correlates early patterns of gene expression with the organization of cells and tissues in the mature, normal animal. The second product will be a knowledge of whether, and if so how, these fates are altered when the gene of interest is not expressed. They will additionally address the question of compartmentation in the mammalian hindbrain and branchial arches. The maps of normal and mutant cell fates will enormously increase our understanding of the roles played by individual genes in the intricate genetic program that regulates mammalian development, and additionally provide a wealth of new information about cell proliferation, cell mixing, and cell migration in the mammalian embryo. In this manner, the research program will contribute to an improved understanding of normal mammalian development and to the kinds of developmental deficits that arise from alterations in specific gene products.

2/7/12 (Item 1 from file: 286)
0050984 Journal Announcement: 08APR91 Doc Type: 2
Nature, 15 MAR 1991, Vol(No) 251(4999), Page(s) 1351-1355

PRIORITY DATE (DEMAND FOR INTERNATIONAL APPL. FILED PRIOR TO EXPIRATION OF 19TH MONTH FROM PRIORITY DATE)
No of Legal Status: 006

2/7/15 (Item 1 from file: 351)
009204307 WPI Acc No: 92-331739/40
XRAM Acc No: C92-147538
FLP-mediated gene modification in mammalian cells - giving precise modification by recombination and can be used to alter transgenes for therapeutic purposes and analysis of development
Patent Assignee: (SALK) SALK INST BIOLOGICAL STUDIES
Author (Inventor): OGORMAN S V; WAHL G M
Number of Patents: 001
Number of Countries: 016
Patent Family:

CC Number	Kind	Date	Week	
WO 9215694	A1	920917	9240	(Basic)

Priority Data (CC No Date): US 666252 (910308)
Applications (CC,No,Date): WO 92US1899 (920306)
Language: English
EP and/or WO Cited Patents: 10Jnl.Ref; US 4959317; US 4997757
Designated States
(National): CA; JP
(Regional): AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LU; MC; NL; SE
Abstract (Basic): WO 9215694 A

Mammalian recombination system (I) comprises:- (a) FLP recombinase or a nucleotide sequence encoding it, and (b) a first DNA comprising a nucleotide sequence contg. up to 1 FLP recombination target site.

Also new are:- (1) a DNA construct comprising as an autonomous fragment, up to 1 FLP recombination target site, up to 1 restriction endonuclease recognition site, up to 1 marker gene, a bacterial origin of replication and opt. a mammalian cellular or viral origin of replication; (2) a DNA construct like that of (1) where the components are contained as an insert in the FLP recombination target site and a 2nd FLP target site is in tandem with the first; (3) assembling of functional gene(s) which is (are) suitable for activation of expression is mammalian cells. The gene segments, derived from up to 1 gene, are individually inactive but contain up to 1 recombination site and are assembled into a functional DNA by contacting with FLP recombinase; (4) disrupting functional gene expression in mammalian cells. Gene of interest contains up to 1 FLP recombination site and is contacted with FLP recombinase and a DNA segment also contg. up to 1 FLP site; (5) the recovery of transfected DNA from the genome of a transfected organism. The DNA contains a fragment with 2 tandemly oriented FLP recombination sites and is contacted with FLP; (6) precisely targeted integration of DNA into the genome of a host organism. An FLP recombination site is introduced into the genome of compatible cells, DNA contg. a recombination site is integrated using FLP recombinase and transformed cells are then introduced into the subject; (7) a mammalian cell contg. up to 1 FLP recombination site in its geromic DNA; (8) a transgenic, non-human mammal contg. up to 1 FLP recombination site in its genome. (9) analysis of the development of a

Salk Institute scientists have shown that the site specific recombinase enzyme, FLP, from Saccharomyces cerevisiae can be used for gene activation in mammalian cells and have suggested it may be useful to mosaically inactivate or activate transgenes or to efficiently integrate transfected DNA at predetermined chromosomal locations.

2/7/13 (Item 1 from file: 315)
328071 CEABA Accession No.: 24-12-020674 DOCUMENT TYPE: Patent
Title: FLP-mediated gene modification in mammalian cells, and compositions and cells useful therefor.
AUTHOR: Wahl, G. M.; O'Gorman, S. V.
CORPORATE SOURCE: Salk Inst. Biol. StudiesLa Jolla, CA 92037USA
CODEN: PIXXD2
PATENT NUMBER: WO 9215694
PUBLICATION DATE: 17 Sep 1992 (920917) LANGUAGE: English
PRIORITY PATENT APPLICATION(S) & DATE(S): US 666252 (910308)
ABSTRACT: A gene activation/inactivation and site specific integration system which was developed for mammalian cells is disclosed. The system is based on the recombination of transfected sequences by FLP, a recombinase derived from Saccharomyces. FLP was shown to rapidly and precisely recombine copies of its specific target sequence in several cell lines e.g. a chromosomally integrated, silent .beta.-galactosidase reporter gene was activated for expression by FLP-mediated removal of intervening sequences to generate clones of marked cells whilst, the reverse reaction, is used to target transfected DNA to specific chromosomal sites. FLP can therefore mosaically activate or inactivate transgenes for a variety of therapeutic purposes, as well as for analysis of vertebrate development.

2/7/14 (Item 1 from file: 345)
11153531
Legal Status (No,Type,Date,Code,Text)
WO 9215694 P 910308 WO AA PRIORITY (PATENT)
US 666252 A 910308
WO 9215694 P 920306 WO AE APPLICATION DATA (APPL. DATA)
WO 92US1899 A 920306
WO 9215694 P 920917 WO AK DESIGNATED STATES CITED IN A PUBLISHED APPLICATION WITH SEARCH REPORT (DESIGNATED STATES CITED IN A PUBLISHED APPL. WITH SEARCH REPORT)
CA JP
WO 9215694 P 920917 WO AL DESIGNATED COUNTRIES FOR REGIONAL PATENTS CITED IN A PUBLISHED APPLICATION WITH SEARCH REPORT (DESIGNATED COUNTRIES FOR REGIONAL PATENTS CITED IN A PUBLISHED APPL. WITH SEARCH REPORT)
AT BE CH DE DK ES FR GB GR IT LU MC NL SE
WO 9215694 P 920917 WO A1 PUBLICATION OF THE INTERNATIONAL APPLICATION WITH THE INTERNATIONAL SEARCH REPORT (PUB. OF THE INTERNATIONAL APPL. WITH THE INTERNATIONAL SEARCH REPORT)
WO 9215694 P 921223 WO DFPE DEMAND FOR INTERNATIONAL APPLICATION
FILED PRIOR TO EXPIRATION OF 19TH MONTH FROM

mammal comprising:- (a) providing a transgenic mammal comprising:- (i) an expression construct encoding FLP under the control of a condition promoter; and; (ii) a reporter construct under the control of the same or a different promoter. The reporter construct encodes a functional or non-functional gene contg. a recombination site such that functional expression is disrupted or functional expression commences on FLP recombination; and (b) following the developement of the mammal to determine when expression of functional reporter gene product either commences or is disrupted; and (10) as co-transfection assay for the occurrence of FLP-mediated recombination in which the expression construct and reporter construct outlined above are contained within plasmids in a mammalian cell.

USE/ADVANTAGE - (I) allows selective modification of chromosomal or extrachromosomal DNA in mammalian cells. Inheritance of genetic sequences and the fate of genetic sequences during development can be studied in a wide variety of tissues in different organisms. Simple histochemical assays can be used for analysis Dwg.0/3B

Derwent Class: B04; C06; D16;
Int Pat Class: C07H-015/12; C12N-005/16; C12N-015/85

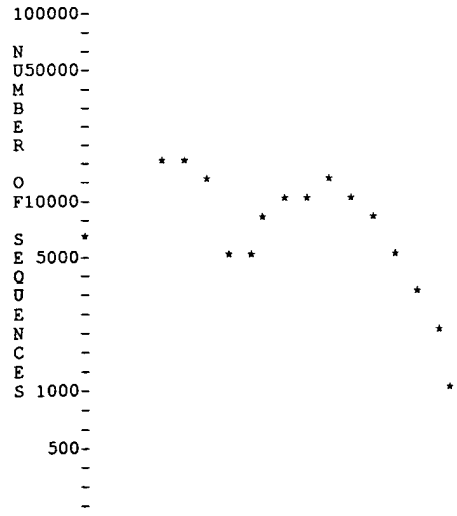
2/7/16 (Item 1 from file: 624)
0432248 DIALOG File 624: McGraw-Hill Publications Online
FLP-mediated gene modification in mammalian cells, and compositions and cells useful therefor
Biotechnology Newswatch November 16, 1992; Pg 9; Vol. 12, No. 22
Journal Code: BIO ISSN: 0275-3687
Section Heading: Biotechnology PatentWatch
Word Count: 143

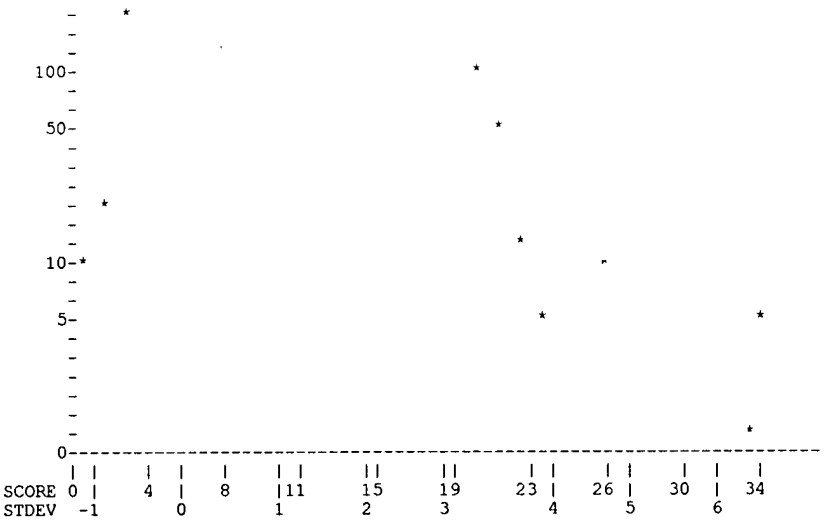
TEXT:
WO 92/15694
Published: Sept. 17, 1992
Filed: March 6, 1992 Priority: March 8, 1991
The Salk Institute For Biological Studies, La Jolla, Ca

A gene activation/inactivation and site-specific integration system has been developed for mammalian cells. The invention system is based on the recombination of transfected sequences by FLP, a recombinase derived from Saccharomyces. In several cell lines, FLP has been shown to rapidly and precisely recombine copies of its specific target sequence. For example, a chromosomally integrated, silent b- galactosidase reporter gene was activated for expression by FLP-mediated removal of intervening sequences to generate clones of marked cells. Alternatively, the reverse reaction can be used to target transfected DNA to specific chromosomal sites. These results demonstrate that FLP can be used, for example, to mosaically activate or inactivate transgenes for a variety of therapeutic purposes, as well as for analysis of vertebrate development.

?

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file flp1.res made by low on Tue 1 Feb 94 14:57:16-PST.

Query sequence being compared:          FLP1 (1-34)
Number of sequences searched:             112413
Number of scores above cutoff:              4909

     Results of the initial comparison of FLP1 (1-34) with:
   Data bank : EMBL-NEW 11, all MAMMALIAN entries
   Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : EMBL-NEW 11, all PRIMATE entries
   Data bank : EMBL-NEW 11, all RODENT entries
   Data bank : GenBank 79, all MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_VERTEBRATE entries
   Data bank : GenBank 79, all PATENT entries
   Data bank : GenBank 79, all PRIMATE entries
   Data bank : GenBank 79, all RODENT entries
   Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
   Data bank : GenBank-NEW 11, all PRIMATE entries
   Data bank : GenBank-NEW 11, all RODENT entries
   Data bank : N-GeneSeq 13, all entries
   Data bank : UEMBL 36_79, all entries
   Data bank : VectorBank 6.4, all entries
```





PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	8	10	4.29

Times:	CPU	Total Elapsed
	00:05:58.96	00:11:52.00

Number of residues: 92888128
Number of sequences searched: 112413
Number of scores above cutoff: 4909

Cut-off raised to 8.
Cut-off raised to 10.
Cut-off raised to 11.
Cut-off raised to 12.
Cut-off raised to 13.
Cut-off raised to 14.
Cut-off raised to 15.
Cut-off raised to 16.

Cut-off raised to 17.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

4 100% similar sequences to the query sequence were found:

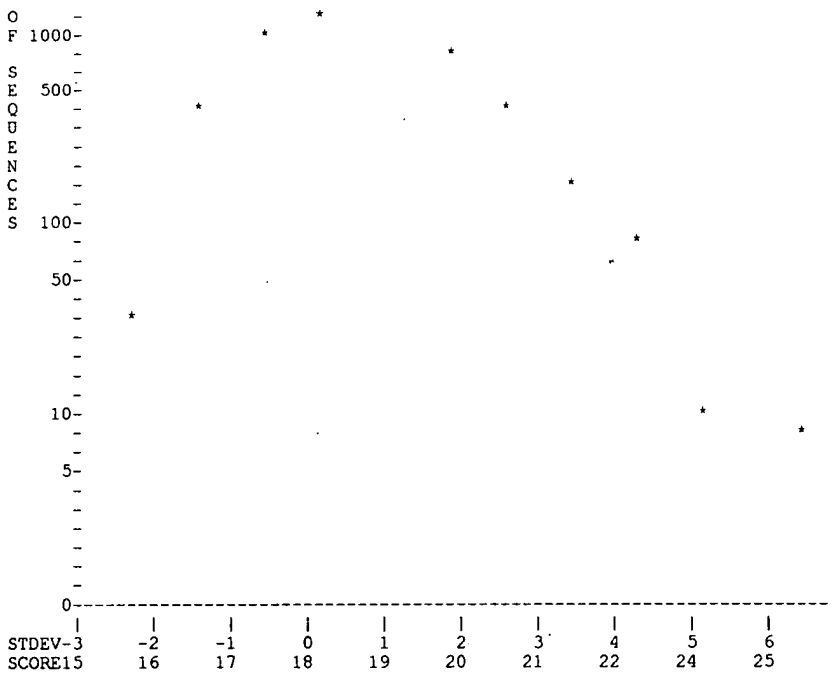
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. Q25185	pSW6 expression vector.	7984	34	34	6.06	0
2. Q44265	pSW6 for expression of LD78 s	7859	34	34	6.06	0
3. Q12154	Shuttle vector pSW6.	7859	34	34	6.06	0
4. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	6.06	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 5 standard deviations above mean ****					
5. Q29100	Sequence of FLP recombination	33	33	33	5.83	0
	**** 3 standard deviations above mean ****					
6. MUSHKPRO	Mouse house-keeping protein m	2415	23	24	3.50	0
7. S56291	sam3=FGFR3 homolog [mice, bra	2520	23	23	3.50	0
8. HSFGL2	Human Flg-2 gene for fibrobla	2887	23	23	3.50	0
9. MUSMFR3	BALB/c fibroblast growth fact	4158	23	23	3.50	0
10. HUMPOMM	Human 35kD peroxisomal membra	1523	22	22	3.27	0

Query sequence being compared: FLP1 (1-34)
Number of sequences optimized: 4909

Results of the optimized comparison of FLP1 (1-34) with:
Data bank : EMBL-NEW 11, all MAMMALIAN entries
Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
Data bank : EMBL-NEW 11, all PRIMATE entries
Data bank : EMBL-NEW 11, all RODENT entries
Data bank : GenBank 79, all MAMMALIAN entries
Data bank : GenBank 79, all OTHER_MAMMALIAN entries
Data bank : GenBank 79, all OTHER_VERTEBRATE entries
Data bank : GenBank 79, all PATENT entries
Data bank : GenBank 79, all PRIMATE entries
Data bank : GenBank 79, all RODENT entries
Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
Data bank : GenBank-NEW 11, all PRIMATE entries
Data bank : GenBank-NEW 11, all RODENT entries
Data bank : N-GeneSeq 13, all entries
Data bank : UEMBL 36_79, all entries
Data bank : VectorBank 6.4, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	18	20	1.23

Times:	CPU	Total Elapsed
	00:01:18.97	00:04:10.00

Number of residues: 14160053
Number of sequences optimized: 4909

The scores below are sorted by optimized score.
Significance is calculated based on optimized score.

4 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. Q25185	pSW6 expression vector.	7984	34	34	13.06	0
2. Q44265	pSW6 for expression of LD78 s	7859	34	34	13.06	0
3. Q12154	Shuttle vector pSW6.	7859	34	34	13.06	0
4. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	13.06	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 12 standard deviations above mean ****					
5. Q29100	Sequence of FLP recombination	33	33	33	12.24	0
	**** 5 standard deviations above mean ****					
6. RSCALPST	Rat mRNA for calpastatin	1931	18	25	5.71	0
7. RABCALPA	Rabbit calpastatin mRNA, comp	3689	22	25	5.71	0
	**** 4 standard deviations above mean ****					
8. MUSHKPRO	Mouse house-keeping protein m	2415	23	24	4.90	0
9. HUMPC1Q1	Human plasma cell membrane gl	3486	16	24	4.90	0
10. Q39050	K.lactis/S. cerevisae genetic	6824	19	24	4.90	0

1. FLP1 (1-34)
Q25185 pSW6 expression vector.

ID Q25185 standard; DNA; 7984 BP.
AC Q25185;
DT 18-NOV-1992 (first entry)
DE pSW6 expression vector.
KW Escherichia coli; 2 micron circle; shuttle vector; leu2; EGF;
KW ampicilin resistant locus; epidermal growth factor; GAL 1-10;
KW phosphoglycerate kinase promoter; PGK; BamHI; HindIII; ss.
OS Saccharomyces cerevisae.
PN WO9207874-A.
PD 14-MAY-1992.
PF 23-OCT-1991; G01860.
PR 24-OCT-1990; GB-023149.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Dawson KM, Edwards RM, Fallon A;
DR WPI; 92-183627/22.
PT New proteins comprising active protein and integrin-affinity
PT sequence - are antithrombotics useful in treating and preventing
PT myocardial infarction, stroke, pulmonary embolism and deep vein
PT thrombosis
PS Disclosure; Page 67; 101pp; English.
CC The sequence given is the yeast expression vector pSW6. It is based
CC on the 2 micron circle from Saccharomyces cerevisae. It is a shuttle
CC vector capable of replication in both S. cerevisae and Escherichia
CC coli as it contains the origin of replication for both organisms. It
CC also contains the leu2 gene (a yeast selectable marker) and the
CC amipicillin resistant locus for selection of plasmid maintenance in E.
CC coli. This vector has enhanced ability for passage through E.coli and
CC this greatly facilitates genetic manipulation with this vector. pSW6
CC contains contains an alpha- factor pre-pro peptide fused in-frame to
CC epidermal growth factor (EGF). The expression of this fusion is under
CC the control of an efficient galactose regulated promoter which contains
CC hybrid DNA sequences from the S. cerevisae GAL 1-10 promoter and the S.
CC cerevisae phosphoglycerate kinase (PGK) promoter. Transcription is
CC terminated in this vector by the natural yeast PGK terminator. The EGF
CC gene in pSW6 can be removed by digestion with HindIII and BamHI. This
CC removes DNA encoding both EGF and 5 amino acids from the C-terminus of
CC the alpha-factor pro-peptide. Genes to be inserted into the pSW6
CC expression vector must therefore have the general composition: HindIII
CC site-alpha-factor adapter-gene-BamHI site.
SQ Sequence 7984 BP; 2348 A; 1698 C; 1635 G; 2303 T;

Initial Score = 34 Optimized Score = 34 Significance = 13.06
Residue Identity = 100% Matches = 34 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```
    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    ||||||||||||||||||||||||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

2. FLP1 (1-34)
Q44265 pSW6 for expression of LD78 synthetic gene.

ID Q44265 standard; DNA; 7859 BP.
AC Q44265;
DT 23-NOV-1993 (first entry)
DE pSW6 for expression of LD78 synthetic gene.
KW SCI; stem cell inhibition; LD78; ACT2; MIP-1alpha;
KW macrophage inflammatory protein; multimer; tumour therapy;
KW psoriasis; hyperproliferation; yeast expression vector;
KW circular; ds.
OS Saccharomyces cerevisiae.
FH Key Location/Qualifiers
FT misc_difference 1773
FT /*tag= a
FT /note= "base illegible in the specification"
PN WO9313206-A.
PD 08-JUL-1993.
PF 23-DEC-1992; G02390.
PR 23-DEC-1991; GB-027319.
PR 14-OCT-1992; GB-021587.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Craig S, Czaplewski LG, Edwards RM, Gilbert RJ;
PI Hunter MG;
DR WPI; 93-227322/28.
PT Protein with stem cell inhibition activity, e.g. LD78 or MIP-1
PT alpha - unable to form stable multimer higher than dodecamer,
PT providing better tissue penetration
PS Disclosure; Page 159-168; 294pp; English.
CC An expression vector was designed to enable secretion of LD78 to
CC the extracellular medium after expression in S. cerevisiae.
CC Secretion aids purification and rapid analysis of LD78.
CC The secretion signals from the yeast mating type factor alpha were
CC used to direct export of the LD78 protein. The yeast expression
CC vector pSW6 (NCIMB 40326) is based on the 2 micron circle from
CC S. cerevisiae.
SQ Sequence 7859 BP; 2317 A; 1667 C; 1585 G; 2289 T;

SQ 0 1 Others;

Initial Score = 34 Optimized Score = 34 Significance = 13.06
Residue Identity = 100% Matches = 34 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```
    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    ||||||||||||||||||||||||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

3. FLP1 (1-34)
Q12154 Shuttle vector pSW6.

```
ID   Q12154 standard; DNA; 7859 BP.
AC   Q12154;
DT   17-SEP-1991  (first entry)
DE   Shuttle vector pSW6.
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; streptokinase; plasmid; circular; ss.
OS   Synthetic.
PN   WO9109125-A.
PD   27-JUN-1991.
PF   07-DEC-1990; G01911.
PR   07-DEC-1989; GB-027722.
PR   07-DEC-1990; WO-G01911.
PA   (BRBI-) BRIT BIO-TECHN LTD.
PI   Dawson KM, Hunter MG, Czapleswski LG;
DR   WPI; 91-208151/28.
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
PS   Disclosure; Page 71; 115pp; English.
CC   The vector is based on the 2u circle from S. cerevisiae.  It is
CC   deposited in S. cerevisiae strain BJ2168 as NCIMB 40326.  It is a
CC   shuttle vector capable of replcation in both E. coli and S. cere-
CC   visiae and contains origins of replication for both, the leu2 gene
CC   (selectable marker), and an ampicillin resistant locus. The E. coli
CC   sequences are derived from E. coli ColE1-based replicon pAT153. The
CC   vector contains an alpha factor pre-pro-peptide gene fused in frame
CC   to the gene for epidermal growth factor (EGF).  The expression of
CC   this fusion is under control of a galactodse regulated promoter
CC   which contains hybrid DNA from S. cerevisiae GAL 1-10 promoter and
CC   the S. cerevisiae phosphoglycerate kinase (PGK) promoter. The EGF
CC   gene can be excised by digestion with HindIII and BamHI. The plas-
CC   mid was used for the expression of a synthetic hirudin HV-1 gene
CC   in E. coli K12 HW87.  The plasmid can be used to construct ex-
CC   pression vectors in which the hirudin gene is linked to a second
CC   gene encoding e.g. another hirudin protein, streptokinase or a
CC   streptokinase-like protein, via a linking peptide.  This peptide
CC   link contains a cleavage site for e.g. factor X or thrombin which
CC   can be cleaved, releasing the individual proteins which have anti-
CC   thrombotic activity.  The enzymes which cleave the fusion protein
CC   are present at the site of the target thrombus so the active agents
CC   are released specifically at the place where clot formation is
CC   occurring.
CC   See also Q12153-Q12156, Q12158-Q12162 and Q12490.
SQ   Sequence  7859 BP;  2317 A;  1656 C;  1600 G;  2286 T;
```

Initial Score = 34 Optimized Score = 34 Significance = 13.06
Residue Identity = 100% Matches = 34 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```
    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    ||||||||||||||||||||||||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

4. FLP1 (1-34)
2MICRON-B B form of the yeast 2micron plasmid.

```
ID   2MICRON-B  standard; DNA; 6248 BP.
XX
AC   IG0001;
XX
DT   09-SEP-1986
XX
DE   B form of the yeast 2micron plasmid.
XX
KW   Vector; circular.
XX
RN   [1]
RA   Broach J.R.;
RT   "The yeast plasmid 2u circle";
RL   Cell 28: 203-204 (1982).
XX
CC   This is the B form of the yeast 2micron plasmid.
CC   Has a single efficient origin of replication that has been
CC   localized to a 350bp site lying largely within one inverted
CC   repeat.  Has two regions of 599bp that are precise inverted
CC   repeats of each other.  Repeats divide the molecule into
CC   approximately equal halves.  There are three ORF, two that
CC   are necessary to maintain the plasmid in high copy number
CC   (REP1 and REP2) and one gene that codes for the FLP protein
CC   responsible for the recombination of the molecule in going
CC   from the A to B forms using the defined protein regions in the
CC   A form in Genbank.  Not available commercially.  No antibiotic
CC   resistance or color markers.
DR   (SUPPLIER (NONE COMMERCIAL))
CC   Key                Location/Qualifiers
CC
CC   pept               3769..2644
CC                       /note="REP1"
CC   pept       4308..5197
CC                       /note="REP2"
CC   pept       5570..6319
CC                       /note="FLP"
CC   repeat_unit       341..938
CC                       /note="inverted repeat"
CC   repeat_unit       3714..4112
CC                       /note="inverted repeat"
CC   orgrpl             700..1050
CC                       /note="2 micron replicon"
SQ   Sequence  6248 BP; 1961 A; 1188 C; 1248 G; 1851 T; 0 other;
```

Initial Score = 34 Optimized Score = 34 Significance = 13.06
Residue Identity = 100% Matches = 34 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```
    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    ||||||||||||||||||||||||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
620       630       640       650  X
```

5. FLP1 (1-34)
   Q29100 Sequence of FLP recombination target site

```
ID   Q29100 standard; DNA; 33 BP.
AC   Q29100;
DT   25-FEB-1992  (first entry)
DE   Sequence of FLP recombination target site
KW   FLP recombinase; site-specific integration system; gene activation;
KW   gene inactivation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    14..21
FT   /*tag= a
FT   /label= spacer
PN   WO9215694-A.
PD   17-SEP-1992.
PF   06-MAR-1992; U01899.
PR   08-MAR-1991; US-666252.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ogorman SV, Wahl GM;
DR   WPI; 92-331739/40.
PT   FLP-mediated gene modification in mammalian cells - giving
PT   precise modification by recombination and can be used to alter
PT   transgenes for therapeutic purposes and analysis of development
PS   Claim 33; Page 40; 49pp; English.
CC   FLP recombinase is a protein which catalyses a site-specific
CC   recombination reaction that is involved in amplifying the copy
CC   number of the 2-mu plasmid of S. cerevisiae during DNA replication.
CC   The inventors claim a mammalian recombination system in which the
CC   FLP recombinase is pref. Q29101. The FLP recombination target site
CC   (FRT) has been identified as minimally comprising two 13 base-pair
CC   repeats, separated by an 8 base-pair spacer (see Q29100). The
CC   nucleotides in the spacer region can be replaced with any other
CC   combination of nucleotides so long as the two 13 base-pair repeats
CC   are separated by 8 nucleotides. NB, in the claims the sequence of
CC   the FRT has only 12 base pairs on the 3' end of the spacer. The
CC   apparently missing base would be C.
SQ   Sequence   33 BP;    11 A;    5 C;    6 G;    11 T;
```

```
Initial Score    =     33  Optimized Score  =     33  Significance = 12.24
Residue Identity =   100%  Matches          =     33  Mismatches   =     0
Gaps             =      0  Conservative Substitutions               =     0
```

```
    X       10        20        30  X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTT
    |||||||||||||||||||||||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTT
    X       10        20        30  X
```

6. FLP1 (1-34)
   RSCALPST Rat mRNA for calpastatin

```
LOCUS       RSCALPST     1931 bp    RNA             ROD       29-MAY-1991
DEFINITION  Rat mRNA for calpastatin
ACCESSION   X56729
KEYWORDS    calpastatin; CANP inhibitor.
SOURCE      rat
  ORGANISM  Rattus sp.
            Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
            Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 1931)
  AUTHORS   Emori,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-1990) Y. Emori, DEPT OF BIOPHYSICS &
            BIOCHEMISTRY, FACULTY OF SCIENCE, UNIVERSITY OF TOKYO, 7-3-1 HONGO,
            BUNKYO-KU, TOKYO 113, JAPAN
  STANDARD  full automatic
REFERENCE   2  (bases 1 to 1931)
  AUTHORS   Ishida,S., Emori,Y. and Suzuki,K.
  TITLE     Rat calpastatin has diverged primary sequence from other mammalian
            calpastatins but retains functionally important sequences
  JOURNAL   Biochim. Biophys. Acta 1088, 436-438 (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     source          1..1931
                     /organism="Rattus sp."
                     /tissue_type="liver"
                     /clone_lib="cDNA"
     mRNA            1..1931
                     /evidence=EXPERIMENTAL
                     /note="calpastatin/CANP inhibitor"
     CDS             18..1829
                     /product="calpastatin/CANP inhibitor"
                     /codon_start=1
                     /translation="MSTTGAKPPVIHEKKPKGKPKEGSETKFQDAPSADGESVAGDVT
                     VATASDEVVVKKKEKKSLTPTLPMESTLNKLSDKSGVNAALDDLIDTLGECEDTNKDD
                     PPYTGPVVLDPMDSTYLEALGIKEGTIPPEYRKLLEKNEAITGPLPDSPKPMGIDHAI
                     DALSSDFTCSSPTGKQTEKEKSTGESSKAQSAGVTRSAVPPQEKKRKVEEEVMNDQAL
                     QALSDSLGTRQPDPQSHLRQAKQVKEAKAKEERQEKCGEDEDTVPAEYRLKPAKDKDG
                     KPLLPEPEETSKCLSESELIGELSADFVQPTYQEKPSMPAAKIKKGVVPDDAVETLAR
                     SLGTRKEDPEDEKSLVDKVKEKAKEEDHEKLGEKEETIPPDYRLEIVKDKDGKPLLPK
                     EAEEQLPPLSDDFLLDALSQDFSSPANILSLGFEDAKLSAAVSETVSQVPAPSNHTAA
                     PPPGTERRDKELDDALDELSDSLGQRQPDPDENKPLDDKVKEKIKAEHSEKLGERDDT
                     IPPEYRHLLDNDGKDKPEKPLDKEHREAGQDQDPIDALSEDLDSCPPTTETSQNTTKE
                     KGKKTSSSKASKNEEKTKDSSKKTEEVPKPKVDEDAT"
BASE COUNT      671 a    406 c    463 g    391 t
ORIGIN
```

```
Initial Score    =     18  Optimized Score  =     25  Significance =  5.71
Residue Identity =    77%  Matches          =     27  Mismatches   =     6
Gaps             =      2  Conservative Substitutions               =     0
```

```
    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATAGG-AACTTC
    ||||   |||||| ||  | ||||||||| ||||||
    GAAGGGACTATTC-CTCCAGAGTATAGGAAACTTC
    X    400       410       420      X
```

7. FLP1 (1-34)
   RABCALPA Rabbit calpastatin mRNA, complete cds.

```
LOCUS       RABCALPA     3689 bp ss-mRNA            MAM       15-JUN-1989
```

```
DEFINITION  Rabbit calpastatin mRNA, complete cds.
ACCESSION   M16476
KEYWORDS    calcium-dependent cysteine protease; calpastatin.
SOURCE      Rabbit lung or heart, cDNA to mRNA, clones
            lambda-C1-[2,311,11,21,213,413,408].
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Lagomorpha; Leporidae.
REFERENCE   1  (bases 1 to 3689)
  AUTHORS   Emori,Y., Kawasaki,H., Imajoh,S., Imahori,K. and Suzuki,K.
  TITLE     Endogenous inhibitor for calcium-dependent cysteine protease
            contains four internal repeats that could be responsible for its
            multiple reactive sites
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 84, 3590-3594 (1987)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     mRNA            <1..3689
                     /note="calpastatin mRNA"
     sig_peptide     160..428
                     /codon_start=1
                     /note="calpastatin signal peptide"
     mat_peptide     429..2313
                     /codon_start=1
                     /note="calpastatin"
     CDS             160..2316
                     /note="calpastatin precursor;  (EC 3.4.22.17)"
                     /codon_start=1
                     /translation="MNPAEAKAVPISKEMEGPHPHSKKRHRRQDAKTEPEKSQSTKPP
                     VDHEKKAQEGKPKEHTKPKSTHKHASDGEGKHGRNEKTASRSKEPVTPAKRTEPETKP
                     QDTKPAGGKSVAAGTTAAPGKAGDPKKEKKSLPAAALAEPKPDEPSGKSGMDAALDDL
                     IDTLGEPSETQEDSTAYTGPEISDPMSSTYIEELGKREVTIPPKYRELLEKKTGVAGP
                     PPDSVTPLGPDDAIDALSSDFTCSSPVASGKEAGKEAAKSAGEVLEAESAKVMRAAAP
                     PQEKKRKVEEDAMSDQALEALSASLGTRMAEPELDLSSIKEVAEAKRKEEKVEKCGED
                     DETVPAEYRLKPATDKDGKPLLPEPAEKPKPRSESELIDELSKDFSQAKSNEKQPKPT
                     GKTEESKAAVPAPVAEAVPRTSMCSIQPVPPKPASLQKSTVPDDAVEALAGSLGRKEA
                     DPEEGKPVADKIKEKSKEEEREKLGEKEETIPPDYRLEEAKDKDGKPLLPSEPTAQLP
                     ALSEDLLLDALSEDFSGPSSASSLKFDDAMLSAAVSEVVSQSPASITRATAPPPDTRP
                     SNKELDDALDKLSDSLGQRQPDPDENKPMEDKVKERAKKEHKDKLGERDDTIPPEYRH
                     LLDQGEQDKPEKPPTKKSKEIKKPAGDQDPIDALSGDLDSCPPAAETSQATEKDKSKT
                     TTASSSKAAKHGDKAKDSAQTTEETSKPKANEKNAS"
     source          1..3689
                     /organism="Oryctolagus cuniculus"
BASE COUNT     1083 a    930 c    943 g    733 t
ORIGIN

Initial Score    =     22  Optimized Score  =     25  Significance =  5.71
Residue Identity =    77%  Matches          =     27  Mismatches   =     6
Gaps             =      2  Conservative Substitutions               =     0

    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATA-GGAACTTC
    |||||| | |||| ||  ||| || | ||||||||
    GAAGTTACAATTC-CTCCAAAATACAGGGAACTTC
    X      760       770       780   X


8. FLP1 (1-34)
   MUSHKPRO     Mouse house-keeping protein mRNA, complete cds.

LOCUS       MUSHKPRO     2415 bp ss-mRNA            ROD       21-AUG-1991
DEFINITION  Mouse house-keeping protein mRNA, complete cds.
ACCESSION   M74555
KEYWORDS    house-keeping protein.
SOURCE      Mus musculus (strain B6) lymphoma cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 2415)
  AUTHORS   Wang,B., Hunsperger,J.P., Laib,J. and Fan,D.
  JOURNAL   Unpublished (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     CDS             88..1278
                     /note="ORF1"
                     /product="house-keeping protein"
                     /codon_start=1
                     /translation="MRGPAMRLPPRLALSALARGPSCILGSGAATRKDWQTRNGRGFS
                     DFNIEPLPDSDLEESSPWTSRNRSEPTRHIACKKAARNLVRDLLEHQNPSRQIILECN
                     PGPGILTGALLKAGARVVAFESEKTFIPHLEPLQRNMDGELQVVHCDFFKMDPRYQEV
                     VRPDVSSQAIFQNLGIKAVPWSAGVPIKVFGILPYKHERRILWKILFDLYSCESIYRY
                     GRVELNMFVSEKEFRKLIATPKRPDLYQVMAVLWQVACDVKFLHMEPWSSFSVHMENG
                     HLEKSKHGESVNLLKQNLYLVRMTPRRTLFTENLSPLNYDIFFHLVKHCFGKRNAPII
                     RHLRSLSTVDPINILRQIRKNPGDTAARMYPHDFKKLFETIEQSEDSVFKWIYDYCPE
                     DMEF"
     source          1..2415
                     /organism="Mus musculus"
BASE COUNT      731 a    478 c    535 g    671 t
ORIGIN

Initial Score    =     23  Optimized Score  =     24  Significance =  4.90
Residue Identity =    70%  Matches          =     24  Mismatches   =    10
Gaps             =      0  Conservative Substitutions               =     0

    X       10        20        30   X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    ||||||| ||| || |  | || | || |||| |
    GAAGTTCATATCCTTTTAACAGAAGAGAAACTAC
    X     1370      1380      1390   X


9. FLP1 (1-34)
   HUMPC1Q1     Human plasma cell membrane glycoprotein (PC-1) mRN

LOCUS       HUMPC1Q1     3486 bp ss-mRNA            PRI       02-NOV-1990
DEFINITION  Human plasma cell membrane glycoprotein (PC-1) mRNA, complete cds.
ACCESSION   M57736 J05654
KEYWORDS    plasma cell membrane glycoprotein PC-1.
SOURCE      Human placenta, cDNA to mRNA, clones lambda-hPC1-2 and
            lambda-hPC1-3; Human fetal liver, cDNA to mRNA, clones
            lambda-hPC1-1 and lambda-hPC1-4.
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 3486)
  AUTHORS   Buckley,M.F., Loveland,K.A., McKinstry,W.J., Garson,O.M. and
            Goding,J.W.
  TITLE     Plasma cell membrane glycoprotein PC-1: cDNA cloning of the human
            molecule, amino acid sequence, and chromosomal location
  JOURNAL   J. Biol. Chem. 265, 17506-17511 (1990)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     mRNA            <1..3238
```

```
                     /gene="PC1"
     mRNA            <1..3484
                     /gene="PC1"
     polyA_signal    3130..3135
                     /gene="PC1"
     CDS             164..2785
                     /gene="PC1"
                     /product="plasma cell membrane glycoprotein PC-1"
                     /codon_start=1
                     /translation="MDVGEEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFG
                     LKPSCAKEVKSCKGRCFERTFGNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRC
                     GEKRLTRSLCACSDDCKDKGDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPT
                     LLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPE
                     SHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVE
                     INGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVS
                     SEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDV
                     KNIKVIYGPAARLRPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFA
                     KSDRIEPLTFYLDPQWQLALNPSERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIE
                     ADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPKEVHPLVQCPFTR
                     NPRDNLGCSCNPSILPIEDFQTQFNLTVAEEKIIKHETLPYGRPRVLQKENTICLLSQ
                     HQFMSGYSQDILMPLWTSYTVDRNDSFSTEDFSNCLYQDFRIPLSPVHKCSFYKNNTK
                     VSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQSFQVIWRYFHDTLLRKYAEERNGVN
                     VVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILIPTHFFIVLTSCKDTSQTPLHCEN
                     LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHRARITDVEHITGLSFYQQRKEPV
                     SDILKLKTHLPTFSQED"
     source          1..3486
                     /organism="Homo sapiens"
BASE COUNT     1022 a     720 c     756 g     988 t
ORIGIN
```

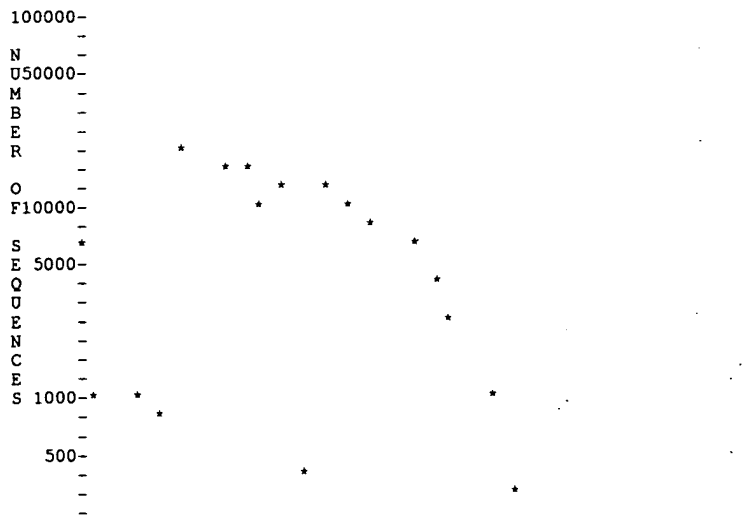
```
Initial Score    =     16  Optimized Score  =     24  Significance =  4.90
Residue Identity =    70%  Matches          =     26  Mismatches   =     8
Gaps             =      3  Conservative Substitutions               =     0

    X         10        20        30   X
    GAA--GTTCCTATTCTCTAGAAAGTATAG-GAACTTC
    |||   ||||| ||| ||  |||||| || ||| ||
    GAATGCTTCCTTTTCACTTAAAAGTAAAGAGAAATTT
  860       870       880       890    X
```

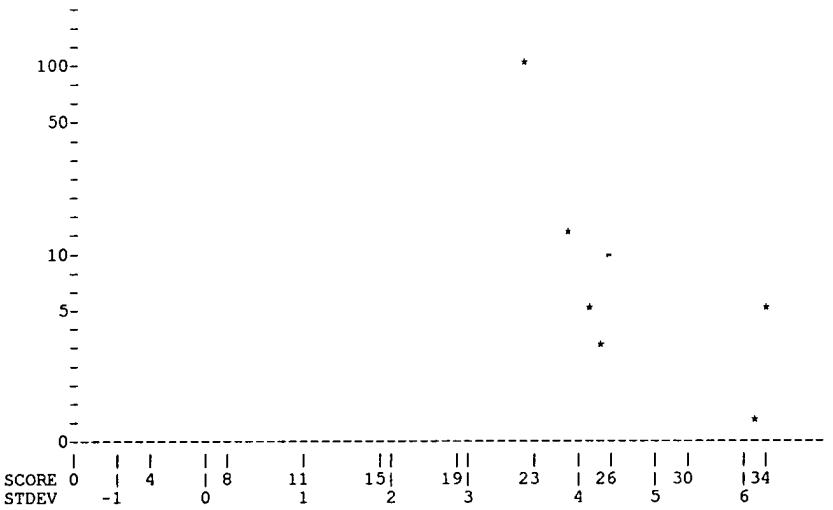
10. FLP1 (1-34)
    Q39050 K.lactis/S. cerevisae genetic vector.

```
ID   Q39050 standard; DNA; 6824 BP.
AC   Q39050;
DT   28-JUL-1993  (first entry)
DE   K.lactis/S. cerevisae genetic vector.
KW   Genetic; vector; integration; Kluyeromyces lactis; 255 ribosomal DNA;
KW   Saccharomyces cerevisiae; E. coli; domain; yeast; plasmid; promoter;
KW   expression cassette; HIS3; marker; transformant; human; lysozyme; HLZ;
KW   GAL7; signal sequence; killer toxin; transcription termination signal;
KW   FLP; 2 micron plasmid; ss.
OS   Synthetic.
PN   EP-537456-A.
PD   21-APR-1993.
PF   31-AUG-1992; 114838.
PR   04-SEP-1991; IT-MI2349.
PA   (ISTS ) SCLAVO SPA.
PI   Galeotti CL,  Gallo E,  Riccio ML,  Rossolini GM,  Thaller MC;
DR   WPI; 93-127394/16.
PT   Vector for Kluyeromyces lactis and Saccharomyces cerevisiae -
PT   which allows stable multiple integration of DNA for prodn. of
PT   heterologous proteins
PS   Claim 1; Fig 1; 26pp; English.
CC   This sequence represents a genetic vector which allows the stable
CC   multiple integration of DNA sequences into the genome of Kluyeromyces
CC   lactis and Saccharomyces cerevisiae.  This sequence can be used in an
CC   integrating vector which comprises a region necessary for the stable
CC   maintenance of the plasmid in E. coli and a domain which acts as an
CC   integrating unit consisting of two not contiguous sequences of the 255
CC   ribosomal DNA from S. cerevisiae, flanking a genetic marker suitable
CC   for selection of the yeast transformants in which the integration
CC   event has occured.  Other DNA sequences may be introduced into the
CC   integration plasmid, such as expression cassettes.  The gene HIS3
CC   from K. lactis and S. cerevisiae is pref. used as a genetic marker
CC   for the selection of transformants and an expression cassette for the
CC   production and secretion into the culture medium of human lysozyme.
CC   This complete transformation vector is 7850 bp long and includes the
CC   integration vector of the invention and an expression cassette
CC   comprising the K. lactis GAL7 promoter, the signal sequence of the K.
CC   lactis killer toxin, the cDNA encoding the ripe form of human lysozyme
CC   (HLZ) and the transcription termination signal FLP of the 2 micron
CC   plasmid from S. cerevisiae.
SQ   Sequence   6824 BP;    1815 A;    1521 C;    1726 G;    1762 T;
```

```
Initial Score    =     19  Optimized Score  =     24  Significance =  4.90
Residue Identity =    65%  Matches          =     25  Mismatches   =     7
Gaps             =      6  Conservative Substitutions               =     0

    X       10        20        30       X
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTT----C
    |||||||||||||||||| ||    |||   |
    GAAGTTCCTATTCTCTAGACAG--CCGGACGGTGGCCA
 2520      2530      2540      2550    X
```

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file flp.res made by low on Tue 1 Feb 94 14:55:48-PST.

Query sequence being compared:          FLP (1-34)
Number of sequences searched:               112413
Number of scores above cutoff:                3751

     Results of the initial comparison of FLP (1-34) with:
   Data bank : EMBL-NEW 11, all MAMMALIAN entries
   Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : EMBL-NEW 11, all PRIMATE entries
   Data bank : EMBL-NEW 11, all RODENT entries
   Data bank : GenBank 79, all MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_VERTEBRATE entries
   Data bank : GenBank 79, all PATENT entries
   Data bank : GenBank 79, all PRIMATE entries
   Data bank : GenBank 79, all RODENT entries
   Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
   Data bank : GenBank-NEW 11, all PRIMATE entries
   Data bank : GenBank-NEW 11, all RODENT entries
   Data bank : N-GeneSeq 13, all entries
   Data bank : UEMBL 36_79, all entries
   Data bank : VectorBank 6.4, all entries
```





PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	9	10	4.34

Times:	CPU	Total Elapsed
	00:05:52.89	00:12:19.00

Number of residues:	92888128
Number of sequences searched:	112413
Number of scores above cutoff:	3751

```
Cut-off raised to 8.
Cut-off raised to 9.
Cut-off raised to 11.
Cut-off raised to 12.
Cut-off raised to 13.
Cut-off raised to 14.
Cut-off raised to 15.
Cut-off raised to 16.
```

Cut-off raised to 17.
Cut-off raised to 18.
Cut-off raised to 19.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

4 100% similar sequences to the query sequence were found:

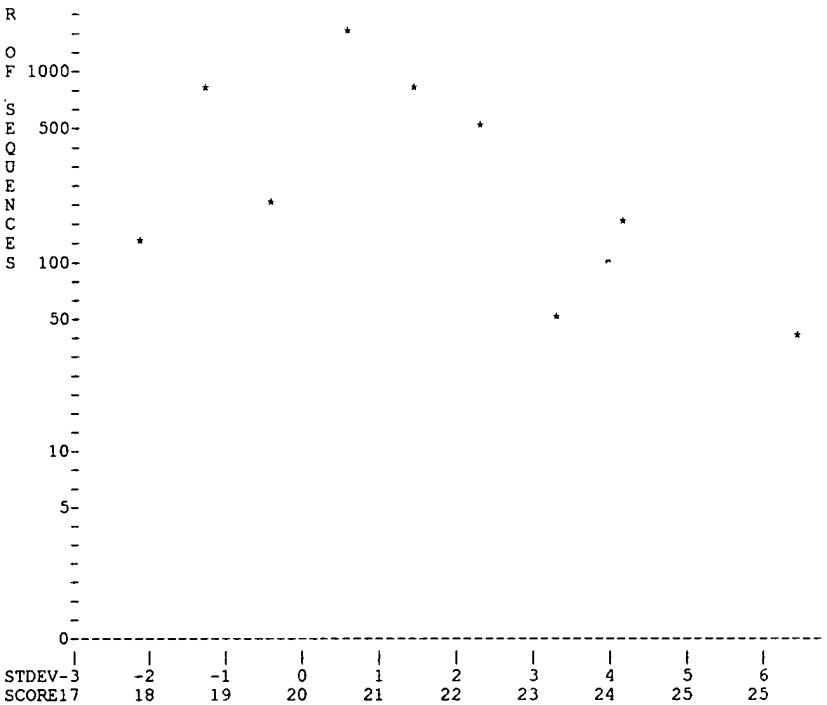
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. Q25185	pSW6 expression vector.	7984	34	34	5.65	0
2. Q44265	pSW6 for expression of LD78 s	7859	34	34	5.65	0
3. Q12154	Shuttle vector pSW6.	7859	34	34	5.65	0
4. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	5.65	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 5 standard deviations above mean ****					
5. Q29100	Sequence of FLP recombination	33	33	33	5.43	0
	**** 3 standard deviations above mean ****					
6. S56291	sam3=FGFR3 homolog [mice, bra	2520	26	26	3.84	0
7. HSFGL2	Human Flg-2 gene for fibrobla	2887	26	26	3.84	0
8. MUSMFR3	BALB/c fibroblast growth fact	4158	26	26	3.84	0
9. PIGCA24TA	Sus scrofa microsatellite pol	211	25	26	3.62	0
10. SMDEXB	S.mutans dextran glucosidase	1800	25	25	3.62	0

Query sequence being compared: FLP (1-34)
Number of sequences optimized: 3751

Results of the optimized comparison of FLP (1-34) with:
Data bank : EMBL-NEW 11, all MAMMALIAN entries
Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
Data bank : EMBL-NEW 11, all PRIMATE entries
Data bank : EMBL-NEW 11, all RODENT entries
Data bank : GenBank 79, all MAMMALIAN entries
Data bank : GenBank 79, all OTHER_MAMMALIAN entries
Data bank : GenBank 79, all OTHER_VERTEBRATE entries
Data bank : GenBank 79, all PATENT entries
Data bank : GenBank 79, all PRIMATE entries
Data bank : GenBank 79, all RODENT entries
Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
Data bank : GenBank-NEW 11, all PRIMATE entries
Data bank : GenBank-NEW 11, all RODENT entries
Data bank : N-GeneSeq_13, all entries
Data bank : UEMBL 36_79, all entries
Data bank : VectorBank 6.4, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	20	21	1.11

Times:	CPU	Total Elapsed
	00:00:56.90	00:03:35.00

Number of residues: 9352894
Number of sequences optimized: 3751

The scores below are sorted by optimized score.
Significance is calculated based on optimized score.

4 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. Q25185	pSW6 expression vector.	7984	34	34	12.64	0
2. Q44265	pSW6 for expression of LD78 s	7859	34	34	12.64	0
3. Q12154	Shuttle vector pSW6.	7859	34	34	12.64	0
4. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	12.64	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 11 standard deviations above mean ****					
5. Q29100	Sequence of FLP recombination	33	33	33	11.73	0
	**** 6 standard deviations above mean ****					
6. RSCALPST	Rat mRNA for calpastatin	1931	22	27	6.32	0
	**** 5 standard deviations above mean ****					
7. HSFGL2	Human Flg-2 gene for fibrobla	2887	26	26	5.42	0
8. MUSMFR3	BALB/c fibroblast growth fact	4158	26	26	5.42	0
9. PIGCA24TA	Sus scrofa microsatellite pol	211	25	26	5.42	0
10. RABCALPA	Rabbit calpastatin mRNA, comp	3689	25	26	5.42	0

1. FLP (1-34)
Q25185 pSW6 expression vector.

ID Q25185 standard; DNA; 7984 BP.
AC Q25185;
DT 18-NOV-1992 (first entry)
DE pSW6 expression vector.
KW Escherichia coli; 2 micron circle; shuttle vector; leu2; EGF;
KW ampicilin resistant locus; epidermal growth factor; GAL 1-10;
KW phosphoglycerate kinase promoter; PGK; BamHI; HindIII; ss.
OS Saccharomyces cerevisae.
PN WO9207874-A.
PD 14-MAY-1992.
PF 23-OCT-1991; G01860.
PR 24-OCT-1990; GB-023149.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Dawson KM, Edwards RM, Fallon A;
DR WPI; 92-183627/22.
PT New proteins comprising active protein and integrin-affinity
PT sequence - are antithrombotics useful in treating and preventing
PT myocardial infarction, stroke, pulmonary embolism and deep vein
PT thrombosis
PS Disclosure; Page 67; 101pp; English.
CC The sequence given is the yeast expression vector pSW6. It is based
CC on the 2 micron circle from Saccharomyces cerevisae. It is a shuttle
CC vector capable of replication in both S. cerevisae and Escherichia
CC coli as it contains the origin of replication for both organisms. It
CC also contains the leu2 gene (a yeast selectable marker) and the
CC amipicillin resistant locus for selection of plasmid maintenance in E.
CC coli. This vector has enhanced ability for passage through E.coli and
CC this greatly facilitates genetic manipulation with this vector. pSW6
CC contains contains an alpha- factor pre-pro peptide fused in-frame to
CC epidermal growth factor (EGF). The expression of this fusion is under
CC the control of an efficient galactose regulated promoter which contains
CC hybrid DNA sequences from the S. cerevisae GAL 1-10 promoter and the S.
CC cerevisae phosphoglycerate kinase (PGK) promoter. Transcription is
CC terminated in this vector by the natural yeast PGK terminator. The EGF
CC gene in pSW6 can be removed by digestion with HindIII and BamHI. This
CC removes DNA encoding both EGF and 5 amino acids from the C-terminus of
CC the alpha-factor pro-peptide. Genes to be inserted into the pSW6
CC expression vector must therefore have the general composition: HindIII
CC site-alpha-factor adapter-gene-BamHI site.
SQ Sequence 7984 BP; 2348 A; 1698 C; 1636 G; 2303 T;

Initial Score = 34 Optimized Score = 34 Significance = 12.64
Residue Identity = 76% Matches = 26 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

```
   X       10        20        30   X
   GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
   ||||||||||||        ||||||||||||
   GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
   X     3140      3150      3160   X
```

2. FLP (1-34)
Q44265 pSW6 for expression of LD78 synthetic gene.

ID Q44265 standard; DNA; 7859 BP.
AC Q44265;
DT 23-NOV-1993 (first entry)
DE pSW6 for expression of LD78 synthetic gene.
KW SCI; stem cell inhibition; LD78; ACT2; MIP-1alpha;
KW macrophage inflammatory protein; multimer; tumour therapy;
KW psoriasis; hyperproliferation; yeast expression vector;
KW circular; ds.
OS Saccharomyces cerevisiae.
FH Key Location/Qualifiers
FT misc_difference 1773
FT /*tag= a
FT /note= "base illegible in the specification"
PN WO9313206-A.
PD 08-JUL-1993.
PF 23-DEC-1992; G02390.
PR 23-DEC-1991; GB-027319.
PR 14-OCT-1992; GB-021587.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Craig S, Czaplewski LG, Edwards RM, Gilbert RJ;
PI Hunter MG;
DR WPI; 93-227322/28.
PT Protein with stem cell inhibition activity, e.g. LD78 or MIP-1
PT alpha - unable to form stable multimer higher than dodecamer,
PT providing better tissue penetration
PS Disclosure; Page 159-168; 294pp; English.
CC An expression vector was designed to enable secretion of LD78 to
CC the extracellular medium after expression in S. cerevisiae.
CC Secretion aids purification and rapid analysis of LD78.
CC The secretion signals from the yeast mating type factor alpha were
CC used to direct export of the LD78 protein. The yeast expression
CC vector pSW6 (NCIMB 40326) is based on the 2 micron circle from

```
 CC   S. cerevisiae.
 SQ   Sequence   7859 BP;    2317 A;    1667 C;    1585 G;    2289 T;
 SQ   1 Others;

Initial Score    =     34  Optimized Score  =     34  Significance = 12.64
Residue Identity =    76%  Matches          =     26  Mismatches   =      8
Gaps             =      0  Conservative Substitutions               =      0

    X       10        20        30   X
    GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
    ||||||||||||        ||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

3. FLP (1-34)
   Q12154   Shuttle vector pSW6.

```
 ID   Q12154 standard; DNA; 7859 BP.
 AC   Q12154;
 DT   17-SEP-1991  (first entry)
 DE   Shuttle vector pSW6.
 KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW   antithrombotic; thrombolysis; streptokinase; plasmid; circular; ss.
 OS   Synthetic.
 PN   WO9109125-A.
 PD   27-JUN-1991.
 PF   07-DEC-1990; G01911.
 PR   07-DEC-1989; GB-027722.
 PR   07-DEC-1990; WO-G01911.
 PA   (BRBI-) BRIT BIO-TECHN LTD.
 PI   Dawson KM, Hunter MG, Czapleswski LG;
 DR   WPI; 91-208151/28.
 PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT   fractions having greater antithrombotic activity for therapy and
 PT   prophylaxis.
 PS   Disclosure; Page 71; 115pp; English.
 CC   The vector is based on the 2u circle from S. cerevisiae.  It is
 CC   deposited in S. cerevisiae strain BJ2168 as NCIMB 40326.  It is a
 CC   shuttle vector capable of replcation in both E. coli and S. cere-
 CC   visiae and contains origins of replication for both, the leu2 gene
 CC   (selectable marker), and an ampicillin resistant locus. The E. coli
 CC   sequences are derived from E. coli ColE1-based replicon pAT153. The
 CC   vector contains an alpha factor pre-pro-peptide gene fused in frame
 CC   to the gene for epidermal growth factor (EGF).  The expression of
 CC   this fusion is under control of a galactodse regulated promoter
 CC   which contains hybrid DNA from S. cerevisiae GAL 1-10 promoter and
 CC   the S. cerevisiae phosphoglycerate kinase (PGK) promoter. The EGF
 CC   gene can be excised by digestion with HindIII and BamHI. The plas-
 CC   mid was used for the expression of a synthetic hirudin HV-1 gene
 CC   in E. coli K12 HW87.   The plasmid can be used to construct ex-
 CC   pression vectors in which the hirudin gene is linked to a second
 CC   gene encoding e.g. another hirudin protein, streptokinase or a
 CC   streptokinase-like protein, via a linking peptide.  This peptide
 CC   link contains a cleavage site for e.g. factor X or thrombin which
 CC   can be cleaved, releasing the individual proteins which have anti-
 CC   thrombotic activity.  The enzymes which cleave the fusion protein
 CC   are present at the site of the target thrombus so the active agents
 CC   are released specifically at the place where clot formation is
 CC   occurring.
 CC   See also Q12153-Q12156, Q12158-Q12162 and Q12490.
 SQ   Sequence   7859 BP;    2317 A;    1656 C;    1600 G;    2286 T;

Initial Score    =     34  Optimized Score  =     34  Significance = 12.64
Residue Identity =    76%  Matches          =     26  Mismatches   =      8
Gaps             =      0  Conservative Substitutions               =      0

    X       10        20        30   X
    GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
    ||||||||||||        ||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

4. FLP (1-34)
   2MICRON-B   B form of the yeast 2micron plasmid.

```
 ID   2MICRON-B  standard; DNA; 6248 BP.
 XX
 AC   IG0001;
 XX
 DT   09-SEP-1986
 XX
 DE   B form of the yeast 2micron plasmid.
 XX
 KW   Vector; circular.
 XX
 RN   [1]
 RA   Broach J.R.;
 RT   "The yeast plasmid 2u circle";
 RL   Cell 28: 203-204 (1982).
 XX
 CC   This is the B form of the yeast 2micron plasmid.
 CC   Has a single efficient origin of replication that has been
 CC   localized to a 350bp site lying largely within one inverted
 CC   repeat.  Has two regions of 599bp that are precise inverted
 CC   repeats of each other.  Repeats divide the molecule into
 CC   approximately equal halves.  There are three ORF, two that
 CC   are necessary to maintain the plasmid in high copy number
 CC   (REP1 and REP2) and one gene that codes for the FLP protein
 CC   responsible for the recombination of the molecule in going
 CC   from the A to B forms using the defined protein regions in the
 CC   A form in Genbank.  Not available commercially.  No antibiotic
 CC   resistance or color markers.
 DR   (SUPPLIER (NONE COMMERCIAL))
 CC   Key               Location/Qualifiers
 CC
 CC   pept              3769..2644
 CC                     /note="REP1"
 CC   pept      4308..5197
 CC                     /note="REP2"
 CC   pept      5570..6319
 CC                     /note="FLP"
 CC   repeat_unit      341..938
 CC                     /note="inverted repeat"
 CC   repeat_unit      3714..4112
 CC                     /note="inverted repeat"
 CC   orgrpl           700..1050
 CC                     /note="2 micron replicon"
 SQ   Sequence  6248 BP; 1961 A; 1188 C; 1248 G; 1851 T; 0 other;

Initial Score    =     34  Optimized Score  =     34  Significance = 12.64
```

```
Residue Identity =    76%  Matches          =     26  Mismatches   =     8
Gaps             =     0%  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
    ||||||||||||        ||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
  620       630       640       650  X
```

5. FLP (1-34)
   Q29100 Sequence of FLP recombination target site

```
ID   Q29100 standard; DNA; 33 BP.
AC   Q29100;
DT   25-FEB-1992  (first entry)
DE   Sequence of FLP recombination target site
KW   FLP recombinase; site-specific integration system; gene activation;
KW   gene inactivation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    14..21
FT   /*tag= a
FT   /label= spacer
PN   WO9215694-A.
PD   17-SEP-1992.
PF   06-MAR-1992; U01899.
PR   08-MAR-1991; US-666252.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ogorman SV, Wahl GM;
DR   WPI; 92-331739/40.
PT   FLP-mediated gene modification in mammalian cells - giving
PT   precise modification by recombination and can be used to alter
PT   transgenes for therapeutic purposes and analysis of development
PS   Claim 33; Page 40; 49pp; English.
CC   FLP recombinase is a protein which catalyses a site-specific
CC   recombination reaction that is involved in amplifying the copy
CC   number of the 2-mu plasmid of S. cerevisiae during DNA replication.
CC   The inventors claim a mammalian recombination system in which the
CC   FLP recombinase is pref. Q29101. The FLP recombination target site
CC   (FRT) has been identified as minimally comprising two 13 base-pair
CC   repeats, separated by an 8 base-pair spacer (see Q29100). The
CC   nucleotides in the spacer region can be replaced with any other
CC   combination of nucleotides so long as the two 13 base-pair repeats
CC   are separated by 8 nucleotides. NB, in the claims the sequence of
CC   the FRT has only 12 base pairs on the 3' end of the spacer. The
CC   apparently missing base would be C.
SQ   Sequence   33 BP;    11 A;    5 C;    6 G;    11 T;
```

```
Initial Score    =     33  Optimized Score  =     33  Significance = 11.73
Residue Identity =    75%  Matches          =     25  Mismatches   =     8
Gaps             =     0%  Conservative Substitutions              =     0

    X       10        20        30  X
    GAAGTTCCTATTCNNNNNNNNGTATAGGAACTT
    ||||||||||||        ||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTT
    X       10        20        30  X
```

6. FLP (1-34)

RSCALPST Rat mRNA for calpastatin

```
LOCUS       RSCALPST     1931 bp    RNA             ROD       29-MAY-1991
DEFINITION  Rat mRNA for calpastatin
ACCESSION   X56729
KEYWORDS    calpastatin; CANP inhibitor.
SOURCE      rat
  ORGANISM  Rattus sp.
            Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
            Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 1931)
  AUTHORS   Emori,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-1990) Y. Emori, DEPT OF BIOPHYSICS &
            BIOCHEMISTRY, FACULTY OF SCIENCE, UNIVERSITY OF TOKYO, 7-3-1 HONGO,
            BUNKYO-KU, TOKYO 113, JAPAN
  STANDARD  full automatic
REFERENCE   2  (bases 1 to 1931)
  AUTHORS   Ishida,S., Emori,Y. and Suzuki,K.
  TITLE     Rat calpastatin has diverged primary sequence from other mammalian
            calpastatins but retains functionally important sequences
  JOURNAL   Biochim. Biophys. Acta 1088, 436-438 (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     source          1..1931
                     /organism="Rattus sp."
                     /tissue_type="liver"
                     /clone_lib="cDNA"
     mRNA            1..1931
                     /evidence=EXPERIMENTAL
                     /note="calpastatin/CANP inhibitor"
     CDS             18..1829
                     /product="calpastatin/CANP inhibitor"
                     /codon_start=1
                     /translation="MSTTGAKPPVIHEKKPKGKPKEGSETKFQDAPSADGESVAGDVT
                     VATASDEVVVKKKEKKSLTPTLPMESTLNKLSDKSGVNAALDDLIDTLGECEDTNKDD
                     PPYTGPVVLDPMDSTYLEALGIKEGTIPPEYRKLLEKNEAITGPLPDSPKPMGIDHAI
                     DALSSDFTCSSPTGKQTEKEKSTGESSKAQSAGVTRSAVPPQEKKRKVEEEVMNDQAL
                     QALSDSLGTRQPDPQSHLRQAKQVKEAKAKEERQEKCGEDEDTVPAEYRLKPAKDKDG
                     KPLLPEPEETSKCLSESELIGELSADFVQPTYQEKPSMPAAKIKKGVVPDDAVETLAR
                     SLGTRKEDPEDEKSLVDKVKEKAKEEDHEKLGEKEETIPPDYRLEIVKDKDGKPLLPK
                     EAEEQLPPLSDDFLLDALSQDFSSPANILSLGFEDAKLSAAVSETVSQVPAPSNHTAA
                     PPPGTERRDKELDDALDELSDSLGQRQPDPDENKPLDDKVKEKIKAEHSEKLGERDDT
                     IPPEYRHLLDNDGKDKPEKPLDKEHREAGQDQDPIDALSEDLDSCPPTTETSQNTTKE
                     KGKKTSSSKASKNEEKTKDSSKKTEEVPKPKVDEDAT"
BASE COUNT      671 a    406 c    463 g    391 t
ORIGIN
```

```
Initial Score    =     22  Optimized Score  =     27  Significance =  6.32
Residue Identity =    65%  Matches          =     23  Mismatches   =    10
Gaps             =     2%  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATTCNNNNNNNNGTATAGG-AACTTC
    ||||   ||||||        ||||||| ||||||
    GAAGGGACTATTC-CTCCAGAGTATAGGAAACTTC
    X    400        410       420     X
```

7. FLP (1-34)
   HSFGL2 Human Flg-2 gene for fibroblast growth factor rece

```
LOCUS       HSFGL2       2887 bp    RNA             PRI       14-AUG-1991
DEFINITION  Human Flg-2 gene for fibroblast growth factor receptor
ACCESSION   X58255
KEYWORDS    fibroblast growth factor receptor; Flg-2 gene.
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
            Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 2887)
  AUTHORS   Givol,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1991) D. Givol, Weizmann Institute of Science,
            76100 Rehovot, Israel
  STANDARD  full automatic
REFERENCE   2  (bases 1 to 2887)
  AUTHORS   Avivi,A., Zimmer,Y., Yayon,A., Yarden,Y. and Givol,D.
  TITLE     Flg-2, a new member of the family of fibroblast growth factor
            receptor
  JOURNAL   Oncogene 6, 1089-1092 (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     source          1..2887
                     /organism="Homo sapiens"
                     /tissue_type="skin"
                     /cell_type="keratinocytes"
                     /tissue_lib="keratinocyte cDNA"
     mRNA            1..2887
                     /gene="Flg-2"
                     /evidence=EXPERIMENTAL
     sig_peptide     213..272
                     /note="fibroblast growth factor receptor"
     mat_peptide     273..2612
                     /gene="Flg-2"
                     /product="fibroblast growth factor receptor"
     CDS             213..2615
                     /gene="Flg-2"
                     /product="fibroblast growth factor receptor"
                     /codon_start=1
                     /translation="MVVPACVLVFCVAVVAGATSEPPGPEQRVVRRAAEVPGPEPSQQ
                     EQVAFGSGDTVELSCHPPGGAPTGPTVWAKDGTGLVASHRILVGPQRLQVLNASHEDA
                     GVYSCQHRLTRRVLCHFSVRVTDAPSSGDDEDGEDVAEDTGAPYWTRPERMDKKLLAV
                     PAANTVRFRCPAAGNPTPSISWLKNGKEFRGQHRIGGIKLRHQQWSLVMESVVPSDRG
                     NYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAILGSDVEFHCKVYSDAQ
                     PHIQWLKHVEVNGSKVGPDGTPYDTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCL
                     AGNSIGFSHHSAWLVVLPAEEELMETDEAGSVYAGVLSYGVVFFLFILVVAAVILCRL
                     RSPPKKGLGSPTVHKVSRFPLKRQVSLESNSSMNSNTPLVRIARLSSGEGPVLANVSE
                     LELPADPKWELSRTRLTLGKPLGEGCFGQVVMAEAIGIDKDRTAKPVTVAVKMLKDDA
                     TDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRP
                     PGMDYSFDACRLPEEQLTCKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNV
                     MKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIF
                     TLGGSPYPGIPVEELFKLLKEGHRMDKPASCTHDLYMIMRECWHAVPSQRPTFKQLVE
                     DLDRILTVTSTDEYLDLSVPFEQYSPGGQDTPSSSSSGDDSVFTHDLLPPGPPSNGGP
                     RT"
BASE COUNT      592 a    834 c    891 g    570 t
ORIGIN

Initial Score    =     26  Optimized Score  =     26  Significance =  5.42
Residue Identity =    58%  Matches          =     20  Mismatches   =    14
Gaps             =      0  Conservative Substitutions               =     0
```

```
   X       10        20        30   X
   GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
   |||| |||| |||         || |  ||||||
   GAAGGTCCTGTTCTGGCCAATGTTTCTGAACTTG
   X   1540      1550      1560     X
```

8. FLP (1-34)
   MUSMFR3 BALB/c fibroblast growth factor receptor 3 (mFR3)

```
LOCUS       MUSMFR3      4158 bp ss-mRNA            ROD       03-SEP-1992
DEFINITION  BALB/c fibroblast growth factor receptor 3 (mFR3) mRNA, complete
            cds.
ACCESSION   M81342 M61881
KEYWORDS    fibroblast growth factor receptor 3; transmembrane protein;
            tyrosine kinase.
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) (library:
            Balb/C brain cDNA library in Lamda ZAP, Stratagene, La Jolla, CA)
            brain cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 4158)
  AUTHORS   Ornitz,D.M. and Leder,P.
  TITLE     Ligand specificity and heparin dependence of fibroblast growth
            factor receptors 1 and 3
  JOURNAL   J. Biol. Chem. 267, 16305-16311 (1992)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     sig_peptide     227..286
                     /gene="mFR3"
                     /codon_start=1
     CDS             227..2632
                     /gene="mFR3"
                     /product="fibroblast growth factor receptor 3"
                     /codon_start=1
                     /translation="MVVPACVLVFCVAVVAGATSEPPGPEQRVVRRAAEVPGPEPSQQ
                     EQVAFGSGDTVELSCHPPGGAPTGPTVWAKDGTGLVASHRILVGPQRLQVLNASHEDA
                     GVYSCQHRLTRRVLCHFSVRVTDAPSSGDDEDGEDVAEDTGAPYWTRPERMDKKLLAV
                     PAANTVRFRCPAAGNPTPSISWLKNGKEFRGEHRIGGIKLRHQQWSLVMESVVPSDRG
                     NYTCVVENKFGSIRQTYTLDVLERSPHRPILQAGLPANQTAILGSDVEFHCKVYSDAQ
                     PHIQWLKHVEVNGSKVGPDGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCL
                     AGNSIGFSHHSAWLVVLPAEEELMETDEAGSVYAGVLSYGVVFFLFILVVAAVILCRL
                     RSPPKKGLGSPTVHKVSRFPLKRQVSLESNSSMNSNTPLVRIARLSSGEGPVLANVSE
                     LELPADPKWELSRTRLTLGKPLGEGCFGQVVMAEAIGIDKDRTAKPVTVAVKMLKDDA
                     TDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRP
                     PGMDYSFDACRLPEEQLTCKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNV
                     MKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIF
                     TPGGPSPYPGIPVEELFKLLKEGHRMDKPASCTHDLYMIMRECWHAVPSQRPTFKQLV
                     EDLDRILTVTSTDEYLDLSVPFEQYSPGGQDTPSSSSSGDDSVFTHDLLPPGPPSNGG
                     PRT"
     source          1..4158
                     /organism="Mus musculus"
BASE COUNT      910 a   1107 c   1205 g    935 t      1 others
ORIGIN

Initial Score    =     26  Optimized Score  =     26  Significance =  5.42
Residue Identity =    58%  Matches          =     20  Mismatches   =    14
Gaps             =      0  Conservative Substitutions               =     0
```

```
   X       10        20        30   X
```

```
    GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
    |||| |||| |||         || |  ||||||
    GAAGGTCCTGTTCTGGCCAATGTTTCTGAACTTG
    1550      1560      1570      1580
```

9. FLP (1-34)
   PIGCA24TA   Sus scrofa microsatellite polymorphism DNA sequenc

```
LOCUS       PIGCA24TA     211 bp ds-DNA             MAM       20-OCT-1992
DEFINITION  Sus scrofa microsatellite polymorphism DNA sequence.
ACCESSION   M77779
KEYWORDS    microsatellite DNA; microsatellite polymorphism.
SOURCE      Sus scrofa (library: pBluescript) blood DNA.
  ORGANISM  Sus scrofa
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Artiodactyla; Suiformes; Suidae.
REFERENCE   1  (bases 1 to 211)
  AUTHORS   Wintero,A.K., Fredholm,M. and Thomsen,P.D.
  TITLE     Variable (dG-dT)n - (dC-dA)n sequences in the porcine genome
  JOURNAL   Genomics 12, 281-288 (1992)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     satellite       66..157
                     /note="perfect and compound; polymorphisim"
     source          1..211
                     /organism="Sus scrofa"
BASE COUNT       76 a      48 c      30 g      57 t
ORIGIN
```

```
Initial Score    =     25  Optimized Score  =     26  Significance =  5.42
Residue Identity =    58%  Matches          =     20  Mismatches   =    14
Gaps             =      0  Conservative Substitutions              =     0
```

```
    X       10        20        30   X
    GAAGTTCCTATTCNNNNNNNNGTATAGGAACTTC
    | |||||  ||||        |  | ||||||||
    GCAGTTCTGATTCGACCCCTAGCCTGGGAACTTC
    X      10        20        30    X
```

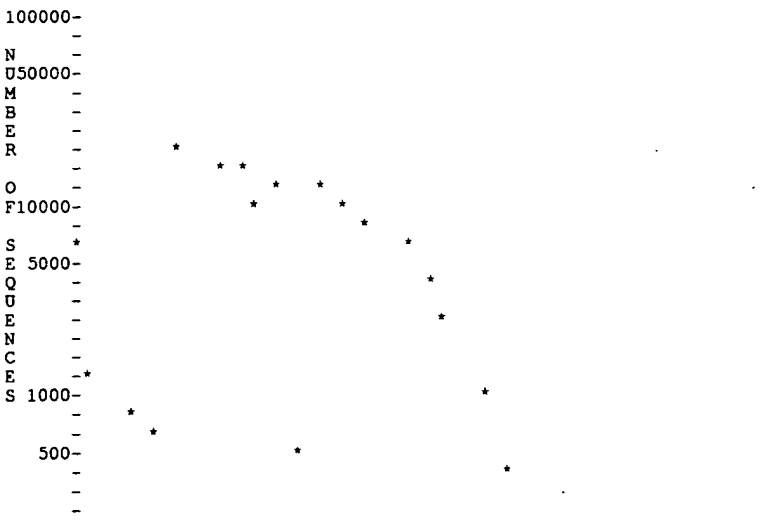
10. FLP (1-34)
    RABCALPA   Rabbit calpastatin mRNA, complete cds.

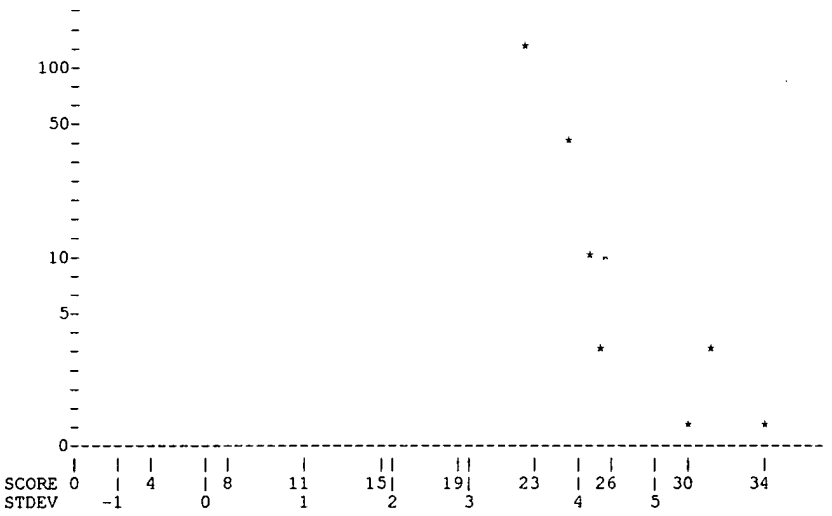
```
LOCUS       RABCALPA      3689 bp ss-mRNA           MAM       15-JUN-1989
DEFINITION  Rabbit calpastatin mRNA, complete cds.
ACCESSION   M16476
KEYWORDS    calcium-dependent cysteine protease; calpastatin.
SOURCE      Rabbit lung or heart, cDNA to mRNA, clones
            lambda-C1-[2,311,11,21,213,413,408].
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Lagomorpha; Leporidae.
REFERENCE   1  (bases 1 to 3689)
  AUTHORS   Emori,Y., Kawasaki,H., Imajoh,S., Imahori,K. and Suzuki,K.
  TITLE     Endogenous inhibitor for calcium-dependent cysteine protease
            contains four internal repeats that could be responsible for its
            multiple reactive sites
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 84, 3590-3594 (1987)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     mRNA            <1..3689
                     /note="calpastatin mRNA"
     sig_peptide     160..428
                     /codon_start=1
                     /note="calpastatin signal peptide"
     mat_peptide     429..2313
                     /codon_start=1
                     /note="calpastatin"
     CDS             160..2316
                     /note="calpastatin precursor;  (EC 3.4.22.17)"
                     /codon_start=1
                     /translation="MNPAEAKAVPISKEMEGPHPHSKKRHRRQDAKTEPEKSQSTKPP
                     VDHEKKAQEGKPKEHTKPKSTHKHASDGEGKHGRNEKTASRSKEPVTPAKRTEPETKP
                     QDTKPAGGKSVAAGTTAAPGKAGDPKKEKKSLPAAALAEPKPDEPSGKSGMDAALDDL
                     IDTLGEPSETQEDSTAYTGPEISDPMSSTYIEELGKREVTIPPKYRELLEKKTGVAGP
                     PPDSVTPLGPDDAIDALSSDFTCSSPVASGKEAGKEAAKSAGEVLEAESAKVMRAAAP
                     PQEKKRKVEEDAMSDQALEALSASLGTRMAEPELDLSSIKEVAEAKRKEEKVEKCGED
                     DETVPAEYRLKPATDKDGKPLLPEPAEKPKPRSESELIDELSKDFSQAKSNEKQPKPT
                     GKTEESKAAVPAPVAEAVPRTSMCSIQPVPPKPASLQKSTVPDDAVEALAGSLGRKEA
                     DPEEGKPVADKIKEKSKEEEREKLGEKEETIPPDYRLEEAKDKDGKPLLPSEPTAQLP
                     ALSEDLLLDALSEDFSGPSSASSLKFDDAMLSAAVSEVVSQSPASITRATAPPPDTRP
                     SNKELDDALDKLSDSLGQRQPDPDENKPMEDKVKERAKKEHKDKLGERDDTIPPEYRH
                     LLDQGEQDKPEKPPTKKSKEIKKPAGDQDPIDALSGDLDSCPPAAETSQATEKDKSKT
                     TTASSSKAAKHGDKAKDSAQTTEETSKPKANEKNAS"
     source          1..3689
                     /organism="Oryctolagus cuniculus"
BASE COUNT     1083 a     930 c     943 g     733 t
ORIGIN
```

```
Initial Score    =     25  Optimized Score  =     26  Significance =  5.42
Residue Identity =    62%  Matches          =     22  Mismatches   =    11
Gaps             =      2  Conservative Substitutions              =     0
```

```
    X       10        20         30   X
    GAAGTTCCTATTCNNNNNNNNGTATA-GGAACTTC
    |||||| | ||||         || | |||||||||
    GAAGTTACAATTC-CTCCAAAATACAGGGAACTTC
    X      760       770        780   X
```

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file flpi.res made by low on Tue 1 Feb 94 15:27:19-PST.

Query sequence being compared:          FLP' (1-34)
Number of sequences searched:               112413
Number of scores above cutoff:                3802

     Results of the initial comparison of FLP' (1-34) with:
   Data bank : EMBL-NEW 11, all MAMMALIAN entries
   Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : EMBL-NEW 11, all PRIMATE entries
   Data bank : EMBL-NEW 11, all RODENT entries
   Data bank : GenBank 79, all MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_VERTEBRATE entries
   Data bank : GenBank 79, all PATENT entries
   Data bank : GenBank 79, all PRIMATE entries
   Data bank : GenBank 79, all RODENT entries
   Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
   Data bank : GenBank-NEW 11, all PRIMATE entries
   Data bank : GenBank-NEW 11, all RODENT entries
   Data bank : N-GeneSeq 13, all entries
   Data bank : UEMBL 36_79, all entries
   Data bank : VectorBank 6.4, all entries
```





PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	9	10	4.39

Times:	CPU	Total Elapsed
	00:06:08.02	00:12:56.00

Number of residues:	92888128
Number of sequences searched:	112413
Number of scores above cutoff:	3802

```
Cut-off raised to 8.
Cut-off raised to 10.
Cut-off raised to 12.
Cut-off raised to 13.
Cut-off raised to 14.
Cut-off raised to 15.
Cut-off raised to 16.
Cut-off raised to 17.
```

Cut-off raised to 18.
Cut-off raised to 19.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

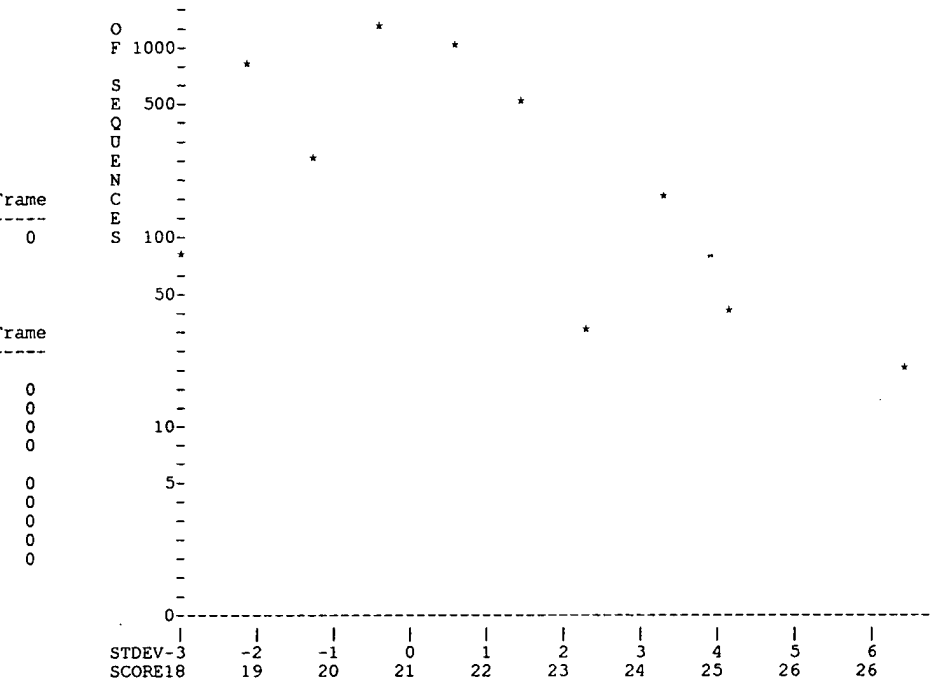
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	5.62	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 4 standard deviations above mean ****					
2. Q44265	pSW6 for expression of LD78 s	7859	31	31	4.95	0
3. Q12154	Shuttle vector pSW6.	7859	31	31	4.95	0
4. Q25185	pSW6 expression vector.	7984	31	31	4.95	0
5. Q29100	Sequence of FLP recombination	33	30	30	4.72	0
	**** 3 standard deviations above mean ****					
6. DOGRAB5A	C.familiaris GTP-binding prot	796	26	26	3.82	0
7. HUMC4BAA	Human complement component C4	848	26	26	3.82	0
8. MUSHKPRO	Mouse house-keeping protein m	2415	26	27	3.82	0
9. SMDEXB	S.mutans dextran glucosidase	1800	25	25	3.60	0
10. RATQLIKEB	Rattus norvegicus Q-like gene	2043	25	26	3.60	0

Query sequence being compared: FLP' (1-34)
Number of sequences optimized: 3802

Results of the optimized comparison of FLP' (1-34) with:
Data bank : EMBL-NEW 11, all MAMMALIAN entries
Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
Data bank : EMBL-NEW 11, all PRIMATE entries
Data bank : EMBL-NEW 11, all RODENT entries
Data bank : GenBank 79, all MAMMALIAN entries
Data bank : GenBank 79, all OTHER_MAMMALIAN entries
Data bank : GenBank 79, all OTHER_VERTEBRATE entries
Data bank : GenBank 79, all PATENT entries
Data bank : GenBank 79, all PRIMATE entries
Data bank : GenBank 79, all RODENT entries
Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
Data bank : GenBank-NEW 11, all PRIMATE entries
Data bank : GenBank-NEW 11, all RODENT entries
Data bank : N-GeneSeq 13, all entries
Data bank : UEMBL 36_79, all entries
Data bank : VectorBank 6.4, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	21	22	1.11

Times:	CPU	Total Elapsed
	00:00:57.94	00:03:32.00

Number of residues: 9177971
Number of sequences optimized: 3802

The scores below are sorted by optimized score.
Significance is calculated based on optimized score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	11.75	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 9 standard deviations above mean ****					
2. Q44265	pSW6 for expression of LD78 s	7859	31	31	9.04	0
3. Q12154	Shuttle vector pSW6.	7859	31	31	9.04	0
4. Q25185	pSW6 expression vector.	7984	31	31	9.04	0
	**** 8 standard deviations above mean ****					
5. Q29100	Sequence of FLP recombination	33	30	30	8.14	0
	**** 5 standard deviations above mean ****					
6. MUSHKPRO	Mouse house-keeping protein m	2415	26	27	5.42	0
	**** 4 standard deviations above mean ****					
7. DOGRAB5A	C.familiaris GTP-binding prot	796	26	26	4.52	0
8. HUMC4BAA	Human complement component C4	848	26	26	4.52	0
9. RATCGM1AC3	Rat carcinoembryonic antigen-	828	18	26	4.52	0
10. RATQLIKEB	Rattus norvegicus Q-like gene	2043	25	26	4.52	0

1. FLP' (1-34)
   2MICRON-B B form of the yeast 2micron plasmid.

```
ID   2MICRON-B  standard; DNA; 6248 BP.
XX
AC   IG0001;
XX
DT   09-SEP-1986
XX
DE   B form of the yeast 2micron plasmid.
XX
KW   Vector; circular.
XX
RN   [1]
RA   Broach J.R.;
RT   "The yeast plasmid 2u circle";
RL   Cell 28: 203-204 (1982).
XX
CC   This is the B form of the yeast 2micron plasmid.
CC   Has a single efficient origin of replication that has been
CC   localized to a 350bp site lying largely within one inverted
CC   repeat.  Has two regions of 599bp that are precise inverted
CC   repeats of each other.  Repeats divide the molecule into
CC   approximately equal halves.  There are three ORF, two that
CC   are necessary to maintain the plasmid in high copy number
CC   (REP1 and REP2) and one gene that codes for the FLP protein
CC   responsible for the recombination of the molecule in going
CC   from the A to B forms using the defined protein regions in the
CC   A form in Genbank.  Not available commercially.  No antibiotic
CC   resistance or color markers.
DR   (SUPPLIER (NONE COMMERCIAL))
CC   Key              Location/Qualifiers
CC
CC   pept             3769..2644
CC                    /note="REP1"
CC   pept       4308..5197
CC                    /note="REP2"
CC   pept       5570..6319
CC                    /note="FLP"
CC   repeat_unit      341..938
CC                    /note="inverted repeat"
CC   repeat_unit      3714..4112
CC                    /note="inverted repeat"
CC   orgrpl           700..1050
CC                    /note="2 micron replicon"
SQ   Sequence  6248 BP; 1961 A; 1188 C; 1248 G; 1851 T; 0 other;
```

Initial Score = 34 Optimized Score = 34 Significance = 11.75
Residue Identity = 76% Matches = 26 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

```
    X       10        20        30   X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
    ||||||||||||        ||||||||||||
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
3860      3870      3880      3890  X
```

2. FLP' (1-34)
   Q44265 pSW6 for expression of LD78 synthetic gene.

```
ID   Q44265 standard; DNA; 7859 BP.
AC   Q44265;
DT   23-NOV-1993  (first entry)
DE   pSW6 for expression of LD78 synthetic gene.
KW   SCI; stem cell inhibition; LD78; ACT2; MIP-1alpha;
KW   macrophage inflammatory protein; multimer; tumour therapy;
KW   psoriasis; hyperproliferation; yeast expression vector;
KW   circular; ds.
OS   Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   misc_difference 1773
FT   /*tag=  a
FT   /note= "base illegible in the specification"
PN   WO9313206-A.
PD   08-JUL-1993.
PF   23-DEC-1992; G02390.
PR   23-DEC-1991; GB-027319.
PR   14-OCT-1992; GB-021587.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI   Craig S,  Czaplewski LG,  Edwards RM,  Gilbert RJ;
PI   Hunter MG;
DR   WPI; 93-227322/28.
PT   Protein with stem cell inhibition activity, e.g. LD78 or MIP-1
PT   alpha - unable to form stable multimer higher than dodecamer,
PT   providing better tissue penetration
PS   Disclosure; Page 159-168; 294pp; English.
CC   An expression vector was designed to enable secretion of LD78 to
CC   the extracellular medium after expression in S. cerevisiae.
CC   Secretion aids purification and rapid analysis of LD78.
```

CC The secretion signals from the yeast mating type factor alpha were
CC used to direct export of the LD78 protein. The yeast expression
CC vector pSW6 (NCIMB 40326) is based on the 2 micron circle from
CC S. cerevisiae.
SQ Sequence 7859 BP; 2317 A; 1667 C; 1585 G; 2289 T;
SQ 1 Others;

Initial Score = 31 Optimized Score = 31 Significance = 9.04
Residue Identity = 70% Matches = 24 Mismatches = 10
Gaps = 0 Conservative Substitutions = 0

```
X       10        20        30   X
GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
|||||||||||| |         | ||||||||||||
GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
X     3140      3150      3160   X
```

3. FLP' (1-34)
Q12154 Shuttle vector pSW6.

ID Q12154 standard; DNA; 7859 BP.
AC Q12154;
DT 17-SEP-1991 (first entry)
DE Shuttle vector pSW6.
KW Fusion protein; blood clotting; coagulation; fibrinolysis;
KW antithrombotic; thrombolysis; streptokinase; plasmid; circular; ss.
OS Synthetic.
PN WO9109125-A.
PD 27-JUN-1991.
PF 07-DEC-1990; G01911.
PR 07-DEC-1989; GB-027722.
PR 07-DEC-1990; WO-G01911.
PA (BRBI-) BRIT BIO-TECHN LTD.
PI Dawson KM, Hunter MG, Czapleswski LG;
DR WPI; 91-208151/28.
PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT fractions having greater antithrombotic activity for therapy and
PT prophylaxis.
PS Disclosure; Page 71; 115pp; English.
CC The vector is based on the 2u circle from S. cerevisiae. It is
CC deposited in S. cerevisiae strain BJ2168 as NCIMB 40326. It is a
CC shuttle vector capable of replcation in both E. coli and S. cere-
CC visiae and contains origins of replication for both, the leu2 gene
CC (selectable marker), and an ampicillin resistant locus. The E. coli
CC sequences are derived from E. coli ColE1-based replicon pAT153. The
CC vector contains an alpha factor pre-pro-peptide gene fused in frame
CC to the gene for epidermal growth factor (EGF). The expression of
CC this fusion is under control of a galactodse regulated promoter
CC which contains hybrid DNA from S. cerevisiae GAL 1-10 promoter and
CC the S. cerevisiae phosphoglycerate kinase (PGK) promoter. The EGF
CC gene can be excised by digestion with HindIII and BamHI. The plas-
CC mid was used for the expression of a synthetic hirudin HV-1 gene
CC in E. coli K12 HW87. The plasmid can be used to construct ex-
CC pression vectors in which the hirudin gene is linked to a second
CC gene encoding e.g. another hirudin protein, streptokinase or a
CC streptokinase-like protein, via a linking peptide. This peptide
CC link contains a cleavage site for e.g. factor X or thrombin which
CC can be cleaved, releasing the individual proteins which have anti-
CC thrombotic activity. The enzymes which cleave the fusion protein
CC are present at the site of the target thrombus so the active agents
CC are released specifically at the place where clot formation is
CC occurring.
CC See also Q12153-Q12156, Q12158-Q12162 and Q12490.
SQ Sequence 7859 BP; 2317 A; 1656 C; 1600 G; 2286 T;

Initial Score = 31 Optimized Score = 31 Significance = 9.04
Residue Identity = 70% Matches = 24 Mismatches = 10
Gaps = 0 Conservative Substitutions = 0

```
X       10        20        30   X
GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
|||||||||||| |         | ||||||||||||
GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
X     3140      3150      3160   X
```

4. FLP' (1-34)
Q25185 pSW6 expression vector.

ID Q25185 standard; DNA; 7984 BP.
AC Q25185;
DT 18-NOV-1992 (first entry)
DE pSW6 expression vector.
KW Escherichia coli; 2 micron circle; shuttle vector; leu2; EGF;
KW ampicilin resistant locus; epidermal growth factor; GAL 1-10;
KW phosphoglycerate kinase promoter; PGK; BamHI; HindIII; ss.
OS Saccharomyces cerevisae.
PN WO9207874-A.
PD 14-MAY-1992.
PF 23-OCT-1991; G01860.
PR 24-OCT-1990; GB-023149.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Dawson KM, Edwards RM, Fallon A;
DR WPI; 92-183627/22.
PT New proteins comprising active protein and integrin-affinity
PT sequence - are antithrombotics useful in treating and preventing
PT myocardial infarction, stroke, pulmonary embolism and deep vein
PT thrombosis
PS Disclosure; Page 67; 101pp; English.
CC The sequence given is the yeast expression vector pSW6. It is based
CC on the 2 micron circle from Saccharomyces cerevisae. It is a shuttle
CC vector capable of replication in both S. cerevisae and Escherichia
CC coli as it contains the origin of replication for both organisms. It
CC also contains the leu2 gene (a yeast selectable marker) and the
CC amipicillin resistant locus for selection of plasmid maintenance in E.
CC coli. This vector has enhanced ability for passage through E.coli and
CC this greatly facilitates genetic manipulation with this vector. pSW6
CC contains contains an alpha- factor pre-pro peptide fused in-frame to
CC epidermal growth factor (EGF). The expression of this fusion is under
CC the control of an efficient galactose regulated promoter which contains
CC hybrid DNA sequences from the S. cerevisae GAL 1-10 promoter and the S.
CC cerevisae phosphoglycerate kinase (PGK) promoter. Transcription is
CC terminated in this vector by the natural yeast PGK terminator. The EGF
CC gene in pSW6 can be removed by digestion with HindIII and BamHI. This
CC removes DNA encoding both EGF and 5 amino acids from the C-terminus of
CC the alpha-factor pro-peptide. Genes to be inserted into the pSW6
CC expression vector must therefore have the general composition: HindIII
CC site-alpha-factor adapter-gene-BamHI site.
SQ Sequence 7984 BP; 2348 A; 1698 C; 1635 G; 2303 T;

Initial Score = 31 Optimized Score = 31 Significance = 9.04

```
Residue Identity =    70%  Matches          =     24  Mismatches   =    10
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
    |||||||||||| |         | ||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

5. FLP' (1-34)
   Q29100 Sequence of FLP recombination target site

```
ID   Q29100 standard; DNA; 33 BP.
AC   Q29100;
DT   25-FEB-1992  (first entry)
DE   Sequence of FLP recombination target site
KW   FLP recombinase; site-specific integration system; gene activation;
KW   gene inactivation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    14..21
FT   /*tag= a
FT   /label= spacer
PN   WO9215694-A.
PD   17-SEP-1992.
PF   06-MAR-1992; 001899.
PR   08-MAR-1991; US-666252.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ogorman SV, Wahl GM;
DR   WPI; 92-331739/40.
PT   FLP-mediated gene modification in mammalian cells - giving
PT   precise modification by recombination and can be used to alter
PT   transgenes for therapeutic purposes and analysis of development
PS   Claim 33; Page 40; 49pp; English.
CC   FLP recombinase is a protein which catalyses a site-specific
CC   recombination reaction that is involved in amplifying the copy
CC   number of the 2-mu plasmid of S. cerevisiae during DNA replication.
CC   The inventors claim a mammalian recombination system in which the
CC   FLP recombinase is pref. Q29101. The FLP recombination target site
CC   (FRT) has been identified as minimally comprising two 13 base-pair
CC   repeats, separated by an 8 base-pair spacer (see Q29100). The
CC   nucleotides in the spacer region can be replaced with any other
CC   combination of nucleotides so long as the two 13 base-pair repeats
CC   are separated by 8 nucleotides. NB, in the claims the sequence of
CC   the FRT has only 12 base pairs on the 3' end of the spacer. The
CC   apparently missing base would be C.
SQ   Sequence   33 BP;    11 A;    5 C;    6 G;    11 T;
```

```
Initial Score    =     30  Optimized Score  =     30  Significance =  8.14
Residue Identity =    69%  Matches          =     23  Mismatches   =    10
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30  X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTT
    |||||||||||| |         | ||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTT
    X       10        20        30  X
```

6. FLP' (1-34)

```
  MUSHKPRO     Mouse house-keeping protein mRNA, complete cds.

LOCUS       MUSHKPRO     2415 bp ss-mRNA            ROD       21-AUG-1991
DEFINITION  Mouse house-keeping protein mRNA, complete cds.
ACCESSION   M74555
KEYWORDS    house-keeping protein.
SOURCE      Mus musculus (strain B6) lymphoma cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 2415)
  AUTHORS   Wang,B., Hunsperger,J.P., Laib,J. and Fan,D.
  JOURNAL   Unpublished (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     CDS             88..1278
                     /note="ORF1"
                     /product="house-keeping protein"
                     /codon_start=1
                     /translation="MRGPAMRLPPRLALSALARGPSCILGSGAATRKDWQTRNGRGFS
                     DFNIEPLPDSDLEESSPWTSRNRSEPTRHIACKKAARNLVRDLLEHQNPSRQIILECN
                     PGPGILTGALLKAGARVVAFESEKTFIPHLEPLQRNMDGELQVVHCDFFKMDPRYQEV
                     VRPDVSSQAIFQNLGIKAVPWSAGVPIKVFGILPYKHERRILWKILFDLYSCESIYRY
                     GRVELNMFVSEKEFRKLIATPKRPDLYQVMAVLWQVACDVKFLHMEPWSSFSVHMENG
                     HLEKSKHGESVNLLKQNLYLVRMTPRRTLFTENLSPLNYDIFFHLVKHCFGKRNAPII
                     RHLRSLSTVDPINILRQIRKNPGDTAARMYPHDFKKLFETIEQSEDSVFKWIYDYCPE
                     DMEF"
     source          1..2415
                     /organism="Mus musculus"
BASE COUNT      731 a    478 c    535 g    671 t
ORIGIN
```

```
Initial Score    =     26  Optimized Score  =     27  Significance =  5.42
Residue Identity =    61%  Matches          =     21  Mismatches   =    13
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
    ||||||| ||| |         ||| || |||| |
    GAAGTTCATATCCTTTTAACAGAAGAGAAACTAC
    X     1370      1380      1390   X
```

7. FLP' (1-34)
   DOGRAB5A C.familiaris GTP-binding protein (rab5) mRNA, comp

```
LOCUS       DOGRAB5A     796 bp ss-mRNA            MAM       15-SEP-1990
DEFINITION  C.familiaris GTP-binding protein (rab5) mRNA, complete cds.
ACCESSION   M35520
KEYWORDS    GTP-binding protein.
SOURCE      C.familiaris (strain Madin-Darby; Cocker spaniel) kidney, cDNA to
            mRNA, clone II.
  ORGANISM  Canis familiaris
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Carnivora; Caniformia; Canoidea; Canidae.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Chavrier,P., Parton,R.G., Hauri,H.P., Simons,K. and Zerial,M.
  TITLE     Localization of low molecular weight GTP binding proteins to
            exocytic and endocytic compartments
  JOURNAL   Cell 62, 317-329 (1990)
  STANDARD  full automatic
```

```
COMMENT     Draft entry and computer-readable sequence for [Cell (1990) In
            press] kindly submitted
            by P.Chavrier, 22-JUN-1990.

            Base-pairs 664 to 711 form a synthetic peptide used to raise
            antibodies.
FEATURES             Location/Qualifiers
     CDS             121..768
                     /note="GTP-binding protein;  (rab5)"
                     /codon_start=1
                     /translation="MANRGATRPNGPNTGNKICQFKLVLLGESAVGKSSLVLRFVKGQ
                     FHEFQESTIGAAFLTQTVCLDDTTVKFEIWDTAGQERYHSLAPMYYRGAQAAIVVYDI
                     TNEESFARAKNWVKELQRQASPNIVIALSGNKADLANKRAVDFQEAQSYADDNSLLFM
                     ETSAKTSMNVNEIFMAIAKKLPKNEPQNPGANSARGRGVDLTEPTQPTRSQCCSN"
     source          1..796
                     /organism="Canis familiaris"
BASE COUNT      267 a    163 c    174 g    192 t
ORIGIN

Initial Score    =     26  Optimized Score  =     26  Significance =  4.52
Residue Identity =    58%  Matches          =     20  Mismatches   =    14
Gaps             =      0  Conservative Substitutions               =     0

    X       10        20        30   X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
    ||||||| | ||          | |||| |||||
    GAAGTTCATTTAAGAGACTGAAATTAGGGACTTC
    X     80        90       100     X
```

```
8. FLP' (1-34)
   HUMC4BAA     Human complement component C4b-binding protein bet

LOCUS       HUMC4BAA      848 bp ss-mRNA            PRI       15-JUN-1990
DEFINITION  Human complement component C4b-binding protein beta-chain (C4BP)
            mRNA, complete cds.
ACCESSION   M29964
KEYWORDS    complement component C4b-binding protein.
SOURCE      Human liver, cDNA to mRNA, clones C1 and A8.
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 848)
  AUTHORS   Hillarp,A. and Dahlback,B.
  TITLE     Cloning of cDNA coding for the beta-chain of human complement
            component C4b-binding protein: Sequence homology with the alpha
            chain
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 87, 1183-1187 (1990)
  STANDARD  full automatic
COMMENT     Draft entry and printed sequence for [1] kindly submitted by
            A.Hillarp, 18-NOV-1989.
FEATURES             Location/Qualifiers
     mRNA            <1..848
                     /gene="C4BPB"
                     /note="C4b-binding protein beta-chain"
     sig_peptide     29..79
                     /gene="C4BPB"
                     /codon_start=1
                     /note="C4b-binding protein beta-chain signal peptide"
     mat_peptide     80..784
                     /gene="C4BPB"
                     /codon_start=1
                     /note="C4b-binding protein beta-chain"
     variation       86..88
                     /gene="C4BPB"
                     /note="gca in clone C1; no codon at 86 in clone A8"
     CDS             29..787
                     /gene="C4BPB"
                     /note="C4b-binding protein beta-chain precursor"
                     /codon_start=1
                     /translation="MFFWCACCLMVAWRVSASDAEHCPELPPVDNSIFVAKEVEGQIL
                     GTYVCIKGYHLVGKKTLFCNASKEWDNTTTECRLGHCPDPVLVNGEFSSSGPVNVSDK
                     ITFMCNDHYILKGSNRSQCLEDHTWAPPFPICKSRDCDPPGNPVHGYFEGNNFTLGST
                     ISYYCEDRYYLVGVQEQQCVDGEWSSALPVCKLIQEAPKPECEKALLAFQESKNLCEA
                     MENFMQQLKESGMTMEELKYSLELKKAELKAKEL"
     source          1..848
                     /organism="Homo sapiens"
BASE COUNT      229 a    174 c    226 g    219 t
ORIGIN

Initial Score    =     26  Optimized Score  =     26  Significance =  4.52
Residue Identity =    58%  Matches          =     20  Mismatches   =    14
Gaps             =      0  Conservative Substitutions               =     0

    X       10        20        30   X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
    |||| ||| | ||        ||  ||| |||||
    GAAGCTCCCAAACCAGAGTGTGAGAAGGCACTTC
    X   620       630       640      X
```

```
9. FLP' (1-34)
   RATCGM1AC3   Rat carcinoembryonic antigen-related protein (CGM1

LOCUS       RATCGM1AC3    828 bp ds-DNA             ROD       15-SEP-1990
DEFINITION  Rat carcinoembryonic antigen-related protein (CGM1) gene, intron B.
ACCESSION   M32478 J05417
KEYWORDS    carcinoembryonic antigen-related protein.
SEGMENT     3 of 8
SOURCE      R.norvegicus (strain Sprague-Dawley) liver DNA, clone
            lambda-rnCGM1-1.
  ORGANISM  Rattus norvegicus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   Rebstock,S., Lucas,K., Thompson,J.A. and Zimmermann,W.A.
  TITLE     cDNA and gene analysis imply a novel structure for a rat
            carcinoembryonic antigen-related protein
  JOURNAL   J. Biol. Chem. 265, 7872-7879 (1990)
  STANDARD  full automatic
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by W.Zimmermann, 02-MAR-1990.
FEATURES             Location/Qualifiers
     intron          <1..>828
                     /note="carcinoembryonic antigen-related protein intron B"
     source          1..828
                     /organism="Rattus norvegicus"
BASE COUNT      193 a    255 c    190 g    190 t
ORIGIN

Initial Score    =     18  Optimized Score  =     26  Significance =  4.52
Residue Identity =    60%  Matches          =     21  Mismatches   =    13
```

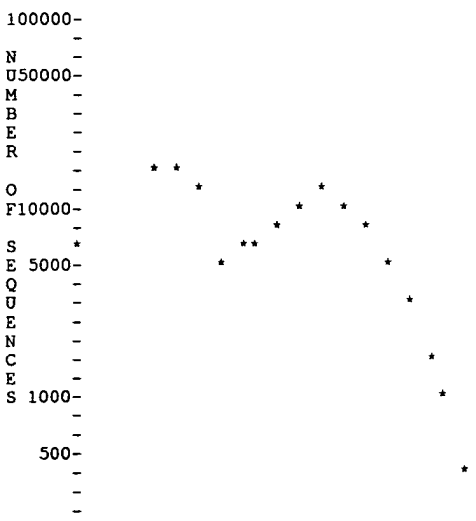
```
Gaps            =      1  Conservative Substitutions             =     0

    X       10        20         30   X
    GAAGTTCCTATACNNNNNNNNGAATAGG-AACTTC
    ||||||||||||          ||  ||| | | ||
    GAAGTTCCTATAAGTGACCAGGAGGAGGCAGCATC
  290       300       310        320  X
```

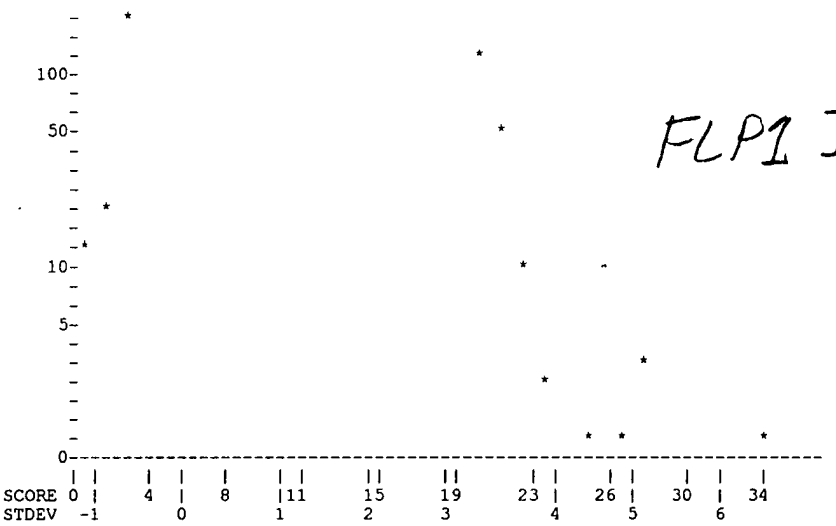
10. FLP' (1-34)
    RATQLIKEB   Rattus norvegicus Q-like gene sequence.

```
LOCUS       RATQLIKEB    2043 bp ds-DNA             ROD       18-MAY-1993
DEFINITION  Rattus norvegicus Q-like gene sequence.
ACCESSION   L16013
KEYWORDS    repeat region.
SOURCE      Rattus norvegicus male adult liver DNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 2043)
  AUTHORS   Rushton,J.J., Misra,D.N., Kunz,H.W., Cortese Hassett,A.L. and
            Gill,T.J.III.
  TITLE     Genomic structure and organization of a Q-like gene in the grc-G/C
            region of the rat
  JOURNAL   Unpublished (1993)
  STANDARD  full automatic
COMMENT     *Locus = RT(2.1).
FEATURES             Location/Qualifiers
     polyA_signal    86..91
                     /note="this polyadenylation signal was selected because of
                     its similarity to RT1.0; putative"
     repeat_region   1013..1264
                     /note="this 252 nucleotide repeat alters the DNA
                     conformation and is associated with at least one rat MHC
                     class I sequence; putative"
     repeat_region   1755..1766
                     /note="ATGC repeat noted but absent in the RT1.0 gene;
                     putative"
     source          1..2043
                     /organism="Rattus norvegicus"
                     /dev_stage="adult"
                     /sex="male"
                     /haplotype="r21"
                     /tissue_type="liver"
                     /sequenced_mol="DNA"
BASE COUNT      387 a    558 c    449 g    649 t
ORIGIN
```

```
Initial Score    =     25  Optimized Score  =     26  Significance =  4.52
Residue Identity =    58%  Matches          =     20  Mismatches   =    14
Gaps             =      0  Conservative Substitutions             =     0

    X       10        20         30   X
    GAAGTTCCTATACNNNNNNNNGAATAGGAACTTC
     | ||||||  ||        ||| | |||||||
    TAGGTTCCTGCACTGGGCTGAGAACATGAACTTC
    X     540       550        560    X
```

```
> o <
o|  |o  IntelliGenetics
> o <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file flp1i.res made by low on Tue 1 Feb 94 15:25:07-PST.


Query sequence being compared:        FLP1' (1-34)
Number of sequences searched:              112413
Number of scores above cutoff:               4804

     Results of the initial comparison of FLP1' (1-34) with:
   Data bank : EMBL-NEW 11, all MAMMALIAN entries
   Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : EMBL-NEW 11, all PRIMATE entries
   Data bank : EMBL-NEW 11, all RODENT entries
   Data bank : GenBank 79, all MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_MAMMALIAN entries
   Data bank : GenBank 79, all OTHER_VERTEBRATE entries
   Data bank : GenBank 79, all PATENT entries
   Data bank : GenBank 79, all PRIMATE entries
   Data bank : GenBank 79, all RODENT entries
   Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
   Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
   Data bank : GenBank-NEW 11, all PRIMATE entries
   Data bank : GenBank-NEW 11, all RODENT entries
   Data bank : N-GeneSeq 13, all entries
   Data bank : UEMBL 36_79, all entries
   Data bank : VectorBank 6.4, all entries
```





FLP1 I.

PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	8	10	4.31

Times:	CPU	Total Elapsed
	00:05:56.98	00:13:12.00

Number of residues: 92888128
Number of sequences searched: 112413
Number of scores above cutoff: 4804

```
Cut-off raised to 8.
Cut-off raised to 10.
Cut-off raised to 11.
Cut-off raised to 12.
Cut-off raised to 13.
Cut-off raised to 14.
Cut-off raised to 15.
Cut-off raised to 16.
```

Cut-off raised to 17.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

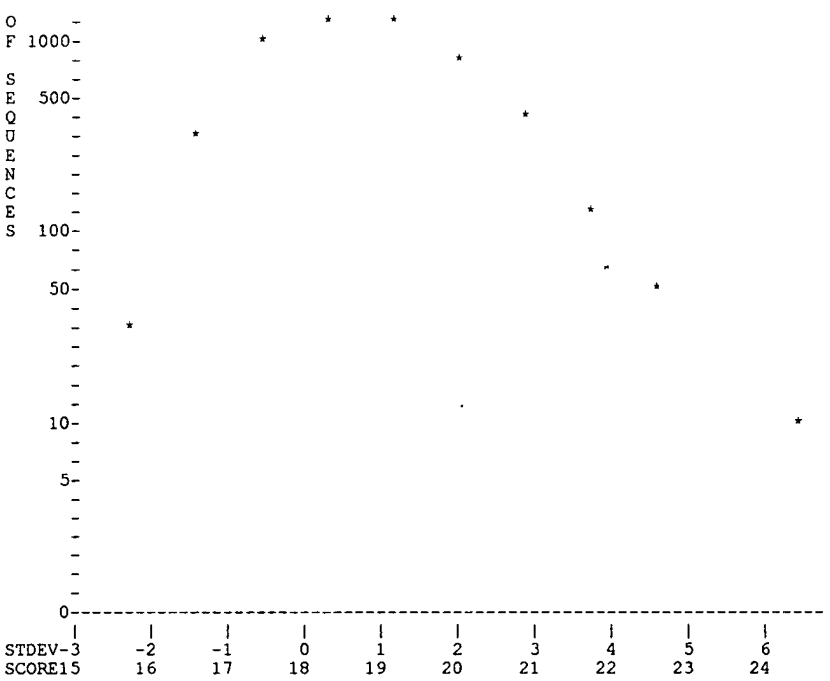
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	6.03	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 4 standard deviations above mean ****					
2. Q44265	pSW6 for expression of LD78 s	7859	28	28	4.64	0
3. Q12154	Shuttle vector pSW6.	7859	28	28	4.64	0
4. Q25185	pSW6 expression vector.	7984	28	28	4.64	0
5. Q29100	Sequence of FLP recombination	33	27	27	4.40	0
	**** 3 standard deviations above mean ****					
6. MUSHKPRO	Mouse house-keeping protein m	2415	25	26	3.94	0
7. RABCALPA	Rabbit calpastatin mRNA, comp	3689	23	23	3.48	0
8. Q23917	Taf DNA polymerase I coding s	4286	23	23	3.48	0
9. DOGRAB5A	C.familiaris GTP-binding prot	796	22	22	3.25	0
10. N82201	Beta-amylase from plant seed.	1794	22	22	3.25	0

Query sequence being compared: FLP1' (1-34)
Number of sequences optimized: 4804

Results of the optimized comparison of FLP1' (1-34) with:
Data bank : EMBL-NEW 11, all MAMMALIAN entries
Data bank : EMBL-NEW 11, all OTHER_MAMMALIAN entries
Data bank : EMBL-NEW 11, all PRIMATE entries
Data bank : EMBL-NEW 11, all RODENT entries
Data bank : GenBank 79, all MAMMALIAN entries
Data bank : GenBank 79, all OTHER_MAMMALIAN entries
Data bank : GenBank 79, all OTHER_VERTEBRATE entries
Data bank : GenBank 79, all PATENT entries
Data bank : GenBank 79, all PRIMATE entries
Data bank : GenBank 79, all RODENT entries
Data bank : GenBank-NEW 11, all OTHER_MAMMALIAN entries
Data bank : GenBank-NEW 11, all OTHER_VERTEBRATE entries
Data bank : GenBank-NEW 11, all PRIMATE entries
Data bank : GenBank-NEW 11, all RODENT entries
Data bank : N-GeneSeq 13, all entries
Data bank : UEMBL 36_79, all entries
Data bank : VectorBank 6.4, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	4
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	1	Number of randomizations	1
Initial scores to save	10	Alignments to save	10
Optimized scores to save	10	Display context	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	18	20	1.19

Times:	CPU	Total Elapsed
	00:01:14.07	00:04:34.00

Number of residues: 13074913
Number of sequences optimized: 4804

The scores below are sorted by optimized score.
Significance is calculated based on optimized score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. 2MICRON-B	B form of the yeast 2micron p	6248	34	34	13.47	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 8 standard deviations above mean ****					
2. Q44265	pSW6 for expression of LD78 s	7859	28	28	8.42	0
3. Q12154	Shuttle vector pSW6.	7859	28	28	8.42	0
4. Q25185	pSW6 expression vector.	7984	28	28	8.42	0
	**** 7 standard deviations above mean ****					
5. Q29100	Sequence of FLP recombination	33	27	27	7.58	0
	**** 6 standard deviations above mean ****					
6. MUSHKPRO	Mouse house-keeping protein m	2415	25	26	6.73	0
	**** 5 standard deviations above mean ****					
7. OCCASB5	Rabbit DNA for 5'flanking reg	2157	19	25	5.89	0
8. RSCALPST	Rat mRNA for calpastatin	1931	19	24	5.05	0
9. HUMH2AZA	Human histone H2A.Z gene, ups	3088	17	24	5.05	0
	**** 4 standard deviations above mean ****					
10. Q23917	Taf DNA polymerase I coding s	4286	23	23	4.21	0

1. FLP1' (1-34)
2MICRON-B B form of the yeast 2micron plasmid.

```
ID   2MICRON-B  standard; DNA; 6248 BP.
XX
AC   IG0001;
XX
DT   09-SEP-1986
XX
DE   B form of the yeast 2micron plasmid.
XX
KW   Vector; circular.
XX
RN   [1]
RA   Broach J.R.;
RT   "The yeast plasmid 2u circle";
RL   Cell 28: 203-204 (1982).
XX
CC   This is the B form of the yeast 2micron plasmid.
CC   Has a single efficient origin of replication that has been
CC   localized to a 350bp site lying largely within one inverted
CC   repeat.  Has two regions of 599bp that are precise inverted
CC   repeats of each other.  Repeats divide the molecule into
CC   approximately equal halves.  There are three ORF, two that
CC   are necessary to maintain the plasmid in high copy number
CC   (REP1 and REP2) and one gene that codes for the FLP protein
CC   responsible for the recombination of the molecule in going
CC   from the A to B forms using the defined protein regions in the
CC   A form in Genbank.  Not available commercially.  No antibiotic
CC   resistance or color markers.
DR   (SUPPLIER (NONE COMMERCIAL))
CC   Key                Location/Qualifiers
CC
CC   pept               3769..2644
CC                       /note="REP1"
CC   pept          4308..5197
CC                       /note="REP2"
CC   pept          5570..6319
CC                       /note="FLP"
CC   repeat_unit        341..938
CC                       /note="inverted repeat"
CC   repeat_unit        3714..4112
CC                       /note="inverted repeat"
CC   orgrpl              700..1050
CC                       /note="2 micron replicon"
SQ   Sequence  6248 BP; 1961 A; 1188 C; 1248 G; 1851 T; 0 other;

Initial Score    =     34  Optimized Score  =     34  Significance = 13.47
Residue Identity =   100%  Matches          =     34  Mismatches   =     0
Gaps             =     0  Conservative Substitutions               =     0

    X       10        20        30   X
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
    |||||||||||||||||||||||||||||||||
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
3860      3870      3880      3890  X
```

2. FLP1' (1-34)
Q44265 pSW6 for expression of LD78 synthetic gene.

```
ID   Q44265 standard; DNA; 7859 BP.
AC   Q44265;
DT   23-NOV-1993  (first entry)
DE   pSW6 for expression of LD78 synthetic gene.
KW   SCI; stem cell inhibition; LD78; ACT2; MIP-1alpha;
KW   macrophage inflammatory protein; multimer; tumour therapy;
KW   psoriasis; hyperproliferation; yeast expression vector;
KW   circular; ds.
OS   Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   misc_difference 1773
FT   /*tag= a
FT   /note= "base illegible in the specification"
PN   WO9313206-A.
PD   08-JUL-1993.
PF   23-DEC-1992; G02390.
PR   23-DEC-1991; GB-027319.
PR   14-OCT-1992; GB-021587.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI   Craig S,  Czaplewski LG,  Edwards RM,  Gilbert RJ;
PI   Hunter MG;
DR   WPI; 93-227322/28.
PT   Protein with stem cell inhibition activity, e.g. LD78 or MIP-1
PT   alpha - unable to form stable multimer higher than dodecamer,
PT   providing better tissue penetration
PS   Disclosure; Page 159-168; 294pp; English.
CC   An expression vector was designed to enable secretion of LD78 to
CC   the extracellular medium after expression in S. cerevisiae.
CC   Secretion aids purification and rapid analysis of LD78.
```

CC The secretion signals from the yeast mating type factor alpha were
CC used to direct export of the LD78 protein. The yeast expression
CC vector pSW6 (NCIMB 40326) is based on the 2 micron circle from
CC S. cerevisiae.
SQ Sequence 7859 BP; 2317 A; 1667 C; 1585 G; 2289 T;
SQ 1 Others;

Initial Score = 28 Optimized Score = 28 Significance = 8.42
Residue Identity = 82% Matches = 28 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

```
X       10        20        30   X
GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
||||||||||| || |  | || ||||||||||||
GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
X     3140      3150      3160   X
```

3. FLP1' (1-34)
Q12154 Shuttle vector pSW6.

ID Q12154 standard; DNA; 7859 BP.
AC Q12154;
DT 17-SEP-1991 (first entry)
DE Shuttle vector pSW6.
KW Fusion protein; blood clotting; coagulation; fibrinolysis;
KW antithrombotic; thrombolysis; streptokinase; plasmid; circular; ss.
OS Synthetic.
PN WO9109125-A.
PD 27-JUN-1991.
PF 07-DEC-1990; G01911.
PR 07-DEC-1989; GB-027722.
PR 07-DEC-1990; WO-G01911.
PA (BRBI-) BRIT BIO-TECHN LTD.
PI Dawson KM, Hunter MG, Czapleswski LG;
DR WPI; 91-208151/28.
PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT fractions having greater antithrombotic activity for therapy and
PT prophylaxis.
PS Disclosure; Page 71; 115pp; English.
CC The vector is based on the 2u circle from S. cerevisiae. It is
CC deposited in S. cerevisiae strain BJ2168 as NCIMB 40326. It is a
CC shuttle vector capable of replcation in both E. coli and S. cere-
CC visiae and contains origins of replication for both, the leu2 gene
CC (selectable marker), and an ampicillin resistant locus. The E. coli
CC sequences are derived from E. coli ColE1-based replicon pAT153. The
CC vector contains an alpha factor pre-pro-peptide gene fused in frame
CC to the gene for epidermal growth factor (EGF). The expression of
CC this fusion is under control of a galactodse regulated promoter
CC which contains hybrid DNA from S. cerevisiae GAL 1-10 promoter and
CC the S. cerevisiae phosphoglycerate kinase (PGK) promoter. The EGF
CC gene can be excised by digestion with HindIII and BamHI. The plas-
CC mid was used for the expression of a synthetic hirudin HV-1 gene
CC in E. coli K12 HW87. The plasmid can be used to construct ex-
CC pression vectors in which the hirudin gene is linked to a second
CC gene encoding e.g. another hirudin protein, streptokinase or a
CC streptokinase-like protein, via a linking peptide. This peptide
CC link contains a cleavage site for e.g. factor X or thrombin which
CC can be cleaved, releasing the individual proteins which have anti-
CC thrombotic activity. The enzymes which cleave the fusion protein
CC are present at the site of the target thrombus so the active agents
CC are released specifically at the place where clot formation is
CC occurring.
CC See also Q12153-Q12156, Q12158-Q12162 and Q12490.
SQ Sequence 7859 BP; 2317 A; 1656 C; 1600 G; 2286 T;

Initial Score = 28 Optimized Score = 28 Significance = 8.42
Residue Identity = 82% Matches = 28 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

```
X       10        20        30   X
GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
||||||||||| || |  | || ||||||||||||
GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
X     3140      3150      3160   X
```

4. FLP1' (1-34)
Q25185 pSW6 expression vector.

ID Q25185 standard; DNA; 7984 BP.
AC Q25185;
DT 18-NOV-1992 (first entry)
DE pSW6 expression vector.
KW Escherichia coli; 2 micron circle; shuttle vector; leu2; EGF;
KW ampicilin resistant locus; epidermal growth factor; GAL 1-10;
KW phosphoglycerate kinase promoter; PGK; BamHI; HindIII; ss.
OS Saccharomyces cerevisae.
PN WO9207874-A.
PD 14-MAY-1992.
PF 23-OCT-1991; G01860.
PR 24-OCT-1990; GB-023149.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI Dawson KM, Edwards RM, Fallon A;
DR WPI; 92-183627/22.
PT New proteins comprising active protein and integrin-affinity
PT sequence - are antithrombotics useful in treating and preventing
PT myocardial infarction, stroke, pulmonary embolism and deep vein
PT thrombosis
PS Disclosure; Page 67; 101pp; English.
CC The sequence given is the yeast expression vector pSW6. It is based
CC on the 2 micron circle from Saccharomyces cerevisae. It is a shuttle
CC vector capable of replication in both S. cerevisae and Escherichia
CC coli as it contains the origin of replication for both organisms. It
CC also contains the leu2 gene (a yeast selectable marker) and the
CC amipicillin resistant locus for selection of plasmid maintenance in E.
CC coli. This vector has enhanced ability for passage through E.coli and
CC this greatly facilitates genetic manipulation with this vector. pSW6
CC contains contains an alpha- factor pre-pro peptide fused in-frame to
CC epidermal growth factor (EGF). The expression of this fusion is under
CC the control of an efficient galactose regulated promoter which contains
CC hybrid DNA sequences from the S. cerevisae GAL 1-10 promoter and the S.
CC cerevisae phosphoglycerate kinase (PGK) promoter. Transcription is
CC terminated in this vector by the natural yeast PGK terminator. The EGF
CC gene in pSW6 can be removed by digestion with HindIII and BamHI. This
CC removes DNA encoding both EGF and 5 amino acids from the C-terminus of
CC the alpha-factor pro-peptide. Genes to be inserted into the pSW6
CC expression vector must therefore have the general composition: HindIII
CC site-alpha-factor adapter-gene-BamHI site.
SQ Sequence 7984 BP; 2348 A; 1698 C; 1635 G; 2303 T;

Initial Score = 28 Optimized Score = 28 Significance = 8.42

```
Residue Identity =    82%  Matches          =     28  Mismatches   =     6
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
    ||||||||||| || |  | || ||||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
    X     3140      3150      3160   X
```

5. FLP1' (1-34)
   Q29100 Sequence of FLP recombination target site

```
ID   Q29100 standard; DNA; 33 BP.
AC   Q29100;
DT   25-FEB-1992  (first entry)
DE   Sequence of FLP recombination target site
KW   FLP recombinase; site-specific integration system; gene activation;
KW   gene inactivation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    14..21
FT   /*tag=  a
FT   /label= spacer
PN   WO9215694-A.
PD   17-SEP-1992.
PF   06-MAR-1992; U01899.
PR   08-MAR-1991; US-666252.
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
PI   Ogorman SV, Wahl GM;
DR   WPI; 92-331739/40.
PT   FLP-mediated gene modification in mammalian cells - giving
PT   precise modification by recombination and can be used to alter
PT   transgenes for therapeutic purposes and analysis of development
PS   Claim 33; Page 40; 49pp; English.
CC   FLP recombinase is a protein which catalyses a site-specific
CC   recombination reaction that is involved in amplifying the copy
CC   number of the 2-mu plasmid of S. cerevisiae during DNA replication.
CC   The inventors claim a mammalian recombination system in which the
CC   FLP recombinase is pref. Q29101. The FLP recombination target site
CC   (FRT) has been identified as minimally comprising two 13 base-pair
CC   repeats, separated by an 8 base-pair spacer (see Q29100). The
CC   nucleotides in the spacer region can be replaced with any other
CC   combination of nucleotides so long as the two 13 base-pair repeats
CC   are separated by 8 nucleotides. NB, in the claims the sequence of
CC   the FRT has only 12 base pairs on the 3' end of the spacer. The
CC   apparently missing base would be C.
SQ   Sequence   33 BP;    11 A;    5 C;    6 G;    11 T;
```

```
Initial Score    =     27  Optimized Score  =     27  Significance =  7.58
Residue Identity =    81%  Matches          =     27  Mismatches   =     6
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30  X
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTT
    ||||||||||| || |  | || ||||||||||
    GAAGTTCCTATTCTCTAGAAAGTATAGGAACTT
    X       10        20        30  X
```

6. FLP1' (1-34)

```
   MUSHKPRO     Mouse house-keeping protein mRNA, complete cds.

LOCUS       MUSHKPRO     2415 bp ss-mRNA            ROD       21-AUG-1991
DEFINITION  Mouse house-keeping protein mRNA, complete cds.
ACCESSION   M74555
KEYWORDS    house-keeping protein.
SOURCE      Mus musculus (strain B6) lymphoma cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 2415)
  AUTHORS   Wang,B., Hunsperger,J.P., Laib,J. and Fan,D.
  JOURNAL   Unpublished (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     CDS             88..1278
                     /note="ORF1"
                     /product="house-keeping protein"
                     /codon_start=1
                     /translation="MRGPAMRLPPRLALSALARGPSCILGSGAATRKDWQTRNGRGFS
                     DFNIEPLPDSDLEESSPWTSRNRSEPTRHIACKKAARNLVRDLLEHQNPSRQIILECN
                     PGPGILTGALLKAGARVVAFESEKTFIPHLEPLQRNMDGELQVVHCDFFKMDPRYQEV
                     VRPDVSSQAIFQNLGIKAVPWSAGVPIKVFGILPYKHERRILWKILFDLYSCESIYRY
                     GRVELNMFVSEKEFRKLIATPKRPDLYQVMAVLWQVACDVKFLHMEPWSSFSVHMENG
                     HLEKSKHGESVNLLKQNLYLVRMTPRRTLFTENLSPLNYDIFFHLVKHCFGKRNAPII
                     RHLRSLSTVDPINILRQIRKNPGDTAARMYPHDFKKLFETIEQSEDSVFKWIYDYCPE
                     DMEF"
     source          1..2415
                     /organism="Mus musculus"
BASE COUNT      731 a    478 c    535 g    671 t
ORIGIN
```

```
Initial Score    =     25  Optimized Score  =     26  Significance =  6.73
Residue Identity =    76%  Matches          =     26  Mismatches   =     8
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
    ||||||| ||| ||||  | |||| || |||| |
    GAAGTTCATATCCTTTTAACAGAAGAGAAACTAC
    X     1370      1380      1390   X
```

7. FLP1' (1-34)
   OCCASB5 Rabbit DNA for 5'flanking region of beta-casein ge

```
LOCUS       OCCASB5      2157 bp    DNA             MAM       23-NOV-1989
DEFINITION  Rabbit DNA for 5'flanking region of beta-casein gene
ACCESSION   X15735
KEYWORDS    beta-casein; hormone-inducible promoter; promoter region.
SOURCE      rabbit
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
            Theria; Eutheria; Lagomorpha; Leporidae.
REFERENCE   1  (bases 1 to 2157)
  AUTHORS   Thomas,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-1989) Thomas B., Swiss Federal Institute of
            Technologie Zuerich, C/O Dr K Buerki Preclinical Research Sandoz
            Ltd, 4002 Basel , Switzerland.
  STANDARD  full automatic
```

```
REFERENCE   2  (bases 1 to 2157)
  AUTHORS   Thomas,B., Bruyere Th. and Buerki,K.
  JOURNAL   Unpublished
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     source          1..2157
                     /organism="Oryctolagus cuniculus"
                     /strain="hare rabbit"
                     /tissue_type="spleen"
                     /clone_lib="lambda-dash"
                     /chromosome="12q24"
     misc_feature    1..2096
                     /note="5'flanking region"
     promoter        2043..2048
                     /note="CAAT-box"
     promoter        2064..2075
                     /note="TATA-box"
     mRNA            2097..2157
                     /note="exon 1"
     misc_feature    2126..2131
                     /note="glucocorticoid receptor-binding site"
BASE COUNT      631 a    461 c    308 g    757 t
ORIGIN

Initial Score    =     19  Optimized Score  =     25  Significance =  5.89
Residue Identity =    74%  Matches          =     26  Mismatches   =     8
Gaps             =      1  Conservative Substitutions               =     0

    X       10        20        30    X
    GAAGTTCCTATACTTTCTAGAGAATAGGAA-CTTC
    ||||| | ||||||  || |||||  || || | ||
    GAAGTGCATATACTCACTTGAGAAAAGAAATCATC
  140       150       160       170  X


8. FLP1' (1-34)
   RSCALPST     Rat mRNA for calpastatin

LOCUS       RSCALPST     1931 bp    RNA             ROD       29-MAY-1991
DEFINITION  Rat mRNA for calpastatin
ACCESSION   X56729
KEYWORDS    calpastatin; CANP inhibitor.
SOURCE      rat
  ORGANISM  Rattus sp.
            Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
            Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 1931)
  AUTHORS   Emori,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-1990) Y. Emori, DEPT OF BIOPHYSICS &
            BIOCHEMISTRY, FACULTY OF SCIENCE, UNIVERSITY OF TOKYO, 7-3-1 HONGO,
            BUNKYO-KU, TOKYO 113, JAPAN
  STANDARD  full automatic
REFERENCE   2  (bases 1 to 1931)
  AUTHORS   Ishida,S., Emori,Y. and Suzuki,K.
  TITLE     Rat calpastatin has diverged primary sequence from other mammalian
            calpastatins but retains functionally important sequences
  JOURNAL   Biochim. Biophys. Acta 1088, 436-438 (1991)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     source          1..1931
                     /organism="Rattus sp."
                     /tissue_type="liver"
                     /clone_lib="cDNA"
     mRNA            1..1931
                     /evidence=EXPERIMENTAL
                     /note="calpastatin/CANP inhibitor"
     CDS             18..1829
                     /product="calpastatin/CANP inhibitor"
                     /codon_start=1
                     /translation="MSTTGAKPPVIHEKKPKGKPKEGSETKFQDAPSADGESVAGDVT
                     VATASDEVVVKKKEKKSLTPTLPMESTLNKLSDKSGVNAALDDLIDTLGECEDTNKDD
                     PPYTGPVVLDPMDSTYLEALGIKEGTIPPEYRKLLEKNEAITGPLPDSPKPMGIDHAI
                     DALSSDFTCSSPTGKQTEKEKSTGESSKAQSAGVTRSAVPPQEKKRKVEEEVMNDQAL
                     QALSDSLGTRQPDPQSHLRQAKQVKEAKAKEERQEKCGEDEDTVPAEYRLKPAKDKDG
                     KPLLPEPEETSKCLSESELIGELSADFVQPTYQEKPSMPAAKIKKGVVPDDAVETLAR
                     SLGTRKEDPEDEKSLVDKVKEKAKEEDHEKLGEKEETIPPDYRLEIVKDKDGKPLLPK
                     EAEEQLPPLSDDFLLDALSQDFSSPANILSLGFEDAKLSAAVSETVSQVPAPSNHTAA
                     PPPGTERRDKELDDALDELSDSLGQRQPDPDENKPLDDKVKEKIKAEHSEKLGERDDT
                     IPPEYRHLLDNDGKDKPEKPLDKEHREAGQDQDPIDALSEDLDSCPPTTETSQNTTKE
                     KGKKTSSSKASKNEEKTKDSSKKTEEVPKPKVDEDAT"
BASE COUNT      671 a    406 c    463 g    391 t
ORIGIN

Initial Score    =     19  Optimized Score  =     24  Significance =  5.05
Residue Identity =    74%  Matches          =     26  Mismatches   =     7
Gaps             =      2  Conservative Substitutions               =     0

    X       10        20        30    X
    GAAGTTCCTATACTTTCTAGAGAATAGG-AACTTC
    ||||   |||| |  || |||| ||||| ||||||
    GAAGGGACTATTC-CTCCAGAGTATAGGAAACTTC
    X    400       410       420     X


9. FLP1' (1-34)
   HUMH2AZA     Human histone H2A.Z gene, upstream promoter sequen

LOCUS       HUMH2AZA     3088 bp ds-DNA             PRI       25-MAY-1993
DEFINITION  Human histone H2A.Z gene, upstream promoter sequence.
ACCESSION   L10137 M33917
KEYWORDS    histone H2A; histone protein.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 3088)
  AUTHORS   Hatch,C.L. and Bonner,W.M.
  TITLE     The human histone H2a.z gene: Sequence and regulation
  JOURNAL   J. Biol. Chem. 265, 15211-15218 (1990)
  STANDARD  full automatic
FEATURES             Location/Qualifiers
     source          1..3088
                     /organism="Homo sapiens"
                     /sequenced_mol="DNA"
BASE COUNT      852 a    687 c    715 g    834 t
ORIGIN

Initial Score    =     17  Optimized Score  =     24  Significance =  5.05
Residue Identity =    73%  Matches          =     25  Mismatches   =     8
Gaps             =      1  Conservative Substitutions               =     0
```

```
    X       10        20        30   X
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
    || | | |||| ||||  | ||||||||||||
    GATGATACTAT-CTTTTCACAGAATAGGAACTCG
    X     1060      1070      1080  X
```

10. FLP1' (1-34)
    Q23917 Taf DNA polymerase I coding sequence.

```
ID   Q23917 standard; DNA; 4286 BP.
AC   Q23917;
DT   27-OCT-1992  (first entry)
DE   Taf DNA polymerase I coding sequence.
KW   Thermostability; PCR; polymerase chain reaction;
KW   thermophilic bacteria; Taf Pol I; ss.
OS   Thermosipho africanus.
FH   Key             Location/Qualifiers
FT   CDS             298..2976
FT   /*tag=  a
FT   /product= Polymerase_I
PN   WO9206202-A.
PD   16-APR-1992.
PF   26-SEP-1991; U07076.
PR   28-SEP-1990; US-590490.
PA   (CETU ) CETUS CORP.
PI   Abramson RD, Gelfand DH, Greenfield L, Lawyer FC, Reichert FL;
DR   WPI; 92-150887/18.
DR   P-PSDB; R23122.
PT   Thermostable DNA polymerase from Thermosipho africanus - prepd.
PT   by purificn. from cells or by expression of Taf polymerase gene
PT   in host cells
PS   Claim 8; Page 6; 80pp; English.
CC   Chromosomal DNA from Thermosipho africanus (Taf) was PCR-amplified
CC   with degenerate primers corresponding to the amino acid sequences
CC   of conserved regions of known thermostable polymerases. When
CC   specific PCR products of a similar size to the product generated
CC   using Taq chromosomal DNA were produced, the PCR fragments were
CC   cloned and sequenced. Fragments with sequences which encoded
CC   regions of amino acid homology to known thermostable polymerases
CC   were identified. The cloned PCR products were used as probes to
CC   screen a genomic Southern blot. The full-length Taf coding sequence
CC   was then compiled from various clones. See also Q23918-Q23961.
SQ   Sequence   4286 BP;    1623 A;    470 C;    847 G;    1346 T;

Initial Score    =     23  Optimized Score  =     23  Significance =  4.21
Residue Identity =    67%  Matches          =     23  Mismatches   =    11
Gaps             =      0  Conservative Substitutions              =     0

    X       10        20        30   X
    GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
    ||||||   | | ||||||  |   ||||||||
    GAAGTTATAAAAGTTTCTAATGTTAAGGAACTTG
    X   3220      3230      3240     X
```

08/148,409

07/666,252

ds

Set	Items	Description
S1	346	FLP(10N)RECOMBINAS?
S2	124	RD (unique items)

?t223/1-124

2/3/1 (Item 1 from file: 155)
08080444 92218444
Reactions between half- and full-FLP recombination target sites. A model system for analyzing early steps in FLP protein-mediated site-specific recombination.
Qian XH; Inman RB; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison 53706.
J Biol Chem (UNITED STATES) Apr 15 1992, 267 (11) p7794-805, ISSN 0021-9258 Journal Code: HIV
Contract/Grant No.: GM-32335; GM-14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/2 (Item 2 from file: 155)
07913378 92051378
FLP-mediated recombination in the vector mosquito, Aedes aegypti.
Morris AC; Schaub TL; James AA
Department of Molecular Biology & Biochemistry, University of California, Irvine 92717.
Nucleic Acids Res (ENGLAND) Nov 11 1991, 19 (21) p5895-900, ISSN 0305-1048 Journal Code: O8L
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/3 (Item 3 from file: 155)
07823652 91342652
Synapsis, strand scission, and strand exchange induced by the FLP recombinase: analysis with half-FRT sites.
Amin A; Roca H; Luetke K; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
Mol Cell Biol Sep 1991, 11 (9) p4497-508, ISSN 0270-7306 Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/4 (Item 4 from file: 155)
07777737 91296737
Domain of a yeast site-specific recombinase (Flp) that recognizes its target site.
Chen JW; Evans BR; Yang SH; Teplow DB; Jayaram M
Department of Microbiology, University of Texas, Austin 78712.
Proc Natl Acad Sci U S A Jul 15 1991, 88 (14) p5944-8, ISSN 0027-8424 Journal Code: PV3
Languages: ENGLISH

Document type: JOURNAL ARTICLE

2/3/5 (Item 5 from file: 155)
07731454 91250454
Identification of the DNA-binding domain of the FLP recombinase.
Pan H; Clary D; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Biol Chem Jun 15 1991, 266 (17) p11347-54, ISSN 0021-9258
Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/6 (Item 6 from file: 155)
07687992 91206992
Integration specificity of retrotransposons and retroviruses.
Sandmeyer SB; Hansen LJ; Chalker DL
Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine 92717.
Annu Rev Genet 1990, 24 p491-518, ISSN 0066-4197 Journal Code: 6DP
Contract/Grant No.: GM33281
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, ACADEMIC

2/3/7 (Item 7 from file: 155)
07668658 91187658
A bacterial model system for chromosomal targeting.
Huang LC; Wood EA; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
Nucleic Acids Res Feb 11 1991, 19 (3) p443-8, ISSN 0305-1048
Journal Code: O8L
Contract/Grant No.: GM37835
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/8 (Item 8 from file: 155)
07645850 91164850
Recombinase-mediated gene activation and site-specific integration in mammalian cells.
O'Gorman S; Fox DT; Wahl GM
Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037.
Science Mar 15 1991, 251 (4999) p1351-5, ISSN 0036-8075
Journal Code: UJ7
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/9 (Item 9 from file: 155)
07643634 91162634
Tyr60 variants of Flp recombinase generate conformationally altered protein-DNA complexes. Differential activity in full-site and half-site recombinations.
Chen JW; Evans BR; Zheng L; Jayaram M
Department of Microbiology, University of Texas at Austin, Austin 78712.

J Mol Biol Mar 5 1991, 218 (1) p107-18, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/10 (Item 10 from file: 155)
07554393 91073393
FLP protein of 2 mu circle plasmid of yeast induces multiple bends in the FLP recognition target site.
Schwartz CJ; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Nov 20 1990, 216 (2) p289-98, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/11 (Item 11 from file: 155)
07553382 91072382
Protein-based asymmetry and protein-protein interactions in FLP recombinase-mediated site-specific recombination.
Qian XH; Inman RB; Cox MM
Program in Cell and Molecular Biology, College of Agricultural and Life Sciences, University of Wisconsin, Madison 53706.
J Biol Chem Dec 15 1990, 265 (35) p21779-88, ISSN 0021-9258
Journal Code: HIV
Contract/Grant No.: GM 37835; GM 14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/12 (Item 12 from file: 155)
07490349 91009349
Identification of the active site tyrosine of Flp recombinase. Possible relevance of its location to the mechanism of recombination [published erratum appears in J Biol Chem 1991 Apr 15;266(11):7312]
Evans BR; Chen JW; Parsons RL; Bauer TK; Teplow DB; Jayaram M
Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037.
J Biol Chem Oct 25 1990, 265 (30) p18504-10, ISSN 0021-9258
Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/13 (Item 13 from file: 155)
07410836 90317836
Synaptic intermediates promoted by the FLP recombinase.
Amin AA; Beatty LG; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Jul 5 1990, 214 (1) p55-72, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/14 (Item 14 from file: 155)
07263960 90170960

Functional analysis of Arg-308 mutants of Flp recombinase. Possible role of Arg-308 in coupling substrate binding to catalysis.
Parsons RL; Evans BR; Zheng L; Jayaram M
Research Institute of Scripps Clinic, La Jolla, California 92037.
J Biol Chem Mar 15 1990, 265 (8) p4527-33, ISSN 0021-9258
Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/15 (Item 15 from file: 155)
07229522 90136522
Use of site-specific recombination to regenerate selectable markers.
Cregg JM; Madden KR
Salk Institute Biotechnology/Industrial Associates, Inc., La Jolla, CA 92037.
Mol Gen Genet Oct 1989, 219 (1-2) p320-3, ISSN 0026-8925
Journal Code: NGP
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/16 (Item 16 from file: 155)
07190832 90097832
Characterization of Holliday structures in FLP protein-promoted site-specific recombination.
Meyer-Leon L; Inman RB; Cox MM
Program in Cellular and Molecular Biology, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706-1569.
Mol Cell Biol Jan 1990, 10 (1) p235-42, ISSN 0270-7306
Journal Code: NGY
Contract/Grant No.: GM37835; GM14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/17 (Item 17 from file: 155)
07123422 90030422
The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome.
Golic KG; Lindquist S
Howard Hughes Medical Institute, Department of Molecular Genetics and Cell Biology, University of Chicago, Illinois 60637.
Cell Nov 3 1989, 59 (3) p499-509, ISSN 0092-8674 Journal Code: CQ4
Contract/Grant No.: GM 25874
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/18 (Item 18 from file: 155)
07011744 89313744
Synthesis of an enzymatically active FLP recombinase in vitro: search for a DNA-binding domain.
Amin AA; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
Mol Cell Biol May 1989, 9 (5) p1987-95, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH

Document type: JOURNAL ARTICLE

2/3/19 (Item 19 from file: 155)
07002130 89304130
FLP-FRT mediated intrachromosomal recombination on a tandemly duplicated YEp integrant at the ILV2 locus of chromosome XIII in Saccharomyces cerevisiae.
Rank GH; Arndt GM; Xiao W
Department of Biology, University of Saskatchewan, Saskatoon, Canada.
Curr Genet Feb 1989, 15 (2) p107-12, ISSN 0172-8083 Journal Code: CUG
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/20 (Item 20 from file: 155)
06876684 89178684
FLP recombinase of the 2 microns circle plasmid of Saccharomyces cerevisiae bends its DNA target. Isolation of FLP mutants defective in DNA bending.
Schwartz CJ; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Feb 20 1989, 205 (4) p647-58, ISSN 0022-2836 Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/21 (Item 21 from file: 155)
06825220 89127220
Holliday intermediates and reaction by-products in FLP protein-promoted site-specific recombination.
Meyer-Leon L; Huang LC; Umlauf SW; Cox MM; Inman RB
Department of Biochemistry, College of Agriculture and Life Sciences, University of Wisconsin-Madison 53706-1569.
Mol Cell Biol Sep 1988, 8 (9) p3784-96, ISSN 0270-7306 Journal Code: NGY
Contract/Grant No.: GM37835; GM14711
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/22 (Item 22 from file: 155)
06823587 89125587
The mechanism of loading of the FLP recombinase onto its DNA target sequence.
Beatty LG; Sadowski PD
Department of Medical Genetics, University of Toronto, Ontario, Canada.
J Mol Biol Nov 20 1988, 204 (2) p283-94, ISSN 0022-2836 Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/23 (Item 23 from file: 155)
06794920 89096920
Step-arrest mutants of FLP recombinase: implications for the catalytic mechanism of DNA recombination.

Parsons RL; Prasad PV; Harshey RM; Jayaram M
Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037.
Mol Cell Biol Aug 1988, 8 (8) p3303-10, ISSN 0270-7306 Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/24 (Item 24 from file: 155)
06761437 89063437
High frequency FLP-independent homologous DNA recombination of 2 mu plasmid in the yeast Saccharomyces cerevisiae.
Bruschi CV; Howe GA
Department of Microbiology and Immunology, School of Medicine, East Carolina University, Greenville, NC 27858-4354.
Curr Genet Sep 1988, 14 (3) p191-9, ISSN 0172-8083 Journal Code: CUG
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/25 (Item 25 from file: 155)
06740094 89042094
Holliday junctions in FLP recombination: resolution by step-arrest mutants of FLP protein.
Jayaram M; Crain KL; Parsons RL; Harshey RM
Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, CA 92037.
Proc Natl Acad Sci U S A Nov 1988, 85 (21) p7902-6, ISSN 0027-8424 Journal Code: PV3
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/26 (Item 26 from file: 155)
06703077 89005077
The functional significance of DNA sequence structure in a site-specific genetic recombination reaction.
Umlauf SW; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
EMBO J Jun 1988, 7 (6) p1845-52, ISSN 0261-4189 Journal Code: EMB
Contract/Grant No.: GM37835; AI00599; GM07215
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/27 (Item 27 from file: 155)
06687975 88332975
DNA recognition by the FLP recombinase of the yeast 2 mu plasmid. A mutational analysis of the FLP binding site.
Senecoff JF; Rossmeissl PJ; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
J Mol Biol May 20 1988, 201 (2) p405-21, ISSN 0022-2836 Journal Code: J6V
Contract/Grant No.: GM37835; AI00599

Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/28 (Item 28 from file: 155)
06643050 88288050
Nucleotide sequencing and expression of the fadL gene involved in long-chain fatty acid transport in Escherichia coli.
Said B; Ghosn CR; Vu L; Nunn WD
Department of Molecular Biology and Biochemistry, University of California, Irvine 92717.
Mol Microbiol May 1988, 2 (3) p363-70, ISSN 0950-382X Journal Code: MOM
Contract/Grant No.: GM 22466-11
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/29 (Item 29 from file: 155)
06618001 88263001
FLP recombinase is an enzyme.
Gates CA; Cox MM
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison 53706.
Proc Natl Acad Sci U S A Jul 1988, 85 (13) p4628-32, ISSN 0027-8424 Journal Code: PV3
Contract/Grant No.: GM37835; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/30 (Item 30 from file: 155)
06567126 88212126
Mutations that improve the binding of yeast FLP recombinase to its substrate.
Lebreton B; Prasad PV; Jayaram M; Youderian P
Department of Biological Sciences, University of Southern California, Los Angeles 90089-1481.
Genetics Mar 1988, 118 (3) p393-400, ISSN 0016-6731 Journal Code: FNH
Contract/Grant No.: GM34982; GM35654
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/31 (Item 31 from file: 155)
06521666 88166666
Antagonistic controls regulate copy number of the yeast 2 mu plasmid.
Murray JA; Scarpa M; Rossi N; Cesareni G
EMBL, Heidelberg, FRG.
EMBO J Dec 20 1987, 6 (13) p4205-12, ISSN 0261-4189 Journal Code: EMB
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/32 (Item 32 from file: 155)
06506025 88151025
Autoregulation of 2 micron circle gene expression provides a model for

maintenance of stable plasmid copy levels.
Som T; Armstrong KA; Volkert FC; Broach JR
Department of Molecular Biology, Princeton University, New Jersey 08544.
Cell Jan 15 1988, 52 (1) p27-37, ISSN 0092-8674 Journal Code: CQ4
Contract/Grant No.: GM34596
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/33 (Item 33 from file: 155)
06342913 87316913
Purification of the FLP site-specific recombinase by affinity chromatography and re-examination of basic properties of the system.
Meyer-Leon L; Gates CA; Attwood JM; Wood EA; Cox MM
Nucleic Acids Res Aug 25 1987, 15 (16) p6469-88, ISSN 0305-1048 Journal Code: O8L
Contract/Grant No.: GM32335; GM37835; AI00599; +
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/34 (Item 34 from file: 155)
06280212 87254212
Isolation of intermediates in the binding of the FLP recombinase of the yeast plasmid 2-micron circle to its target sequence.
Andrews BJ; Beatty LG; Sadowski PD
J Mol Biol Jan 20 1987, 193 (2) p345-58, ISSN 0022-2836 Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/35 (Item 35 from file: 155)
06274060 87248060
Rapid localization and characterization of random mutations within the 2 micron circle site-specific recombinase: a general strategy for analysis of protein function [published erratum appears in Gene 1987;57(1):149]
Govind NS; Jayaram M
Gene 1987, 51 (1) p31-41, ISSN 0378-1119 Journal Code: FOP
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/36 (Item 36 from file: 155)
06210407 87184407
Site-specific recombination of the yeast plasmid two-micron circle: intermediates in the binding process.
Andrews BJ; Beatty LG; Sadowski PD
Basic Life Sci 1986, 40 p407-24, ISSN 0090-5542 Journal Code: 9K0
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/37 (Item 37 from file: 155)
06210406 87184406
Site-specific recombination promoted in vitro by the FLP protein of the yeast two-micron plasmid.
Senecoff JF; Bruckner RC; Meyer-Leon L; Gates CA; Wood E; Umlauf SW; Attwood JM; Cox MM

Basic Life Sci 1986, 40 p397-405, ISSN 0090-5542 Journal Code: 9K0
Contract/Grant No.: GM32335; 5-T32 GM07215; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/38 (Item 38 from file: 155)
06210404 87184404
Survival strategies of the yeast plasmid two-micron circle.
Volkert FC; Wu LC; Fisher PA; Broach JR
Basic Life Sci 1986, 40 p375-96, ISSN 0090-5542 Journal Code: 9K0
Contract/Grant No.: GM34596; GM33132
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/39 (Item 39 from file: 155)
06201639 87175639
Mutations in the 2-microns circle site-specific recombinase that abolish recombination without affecting substrate recognition [published erratum appears in Proc Natl Acad Sci U S A 1988 Mar;85(5):1497]
Prasad PV; Young LJ; Jayaram M
Proc Natl Acad Sci U S A Apr 1987, 84 (8) p2189-93, ISSN 0027-8424
Journal Code: PV3
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/40 (Item 40 from file: 155)
06167165 87141165
Association of reciprocal exchange with gene conversion between the repeated segments of 2-micron circle.
Jayaram M
J Mol Biol Oct 5 1986, 191 (3) p341-54, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/41 (Item 41 from file: 155)
06115790 87089790
Substrate recognition by the 2 micron circle site-specific recombinase: effect of mutations within the symmetry elements of the minimal substrate.
Prasad PV; Horensky D; Young LJ; Jayaram M
Mol Cell Biol Dec 1986, 6 (12) p4329-34, ISSN 0270-7306
Journal Code: NGY
Contract/Grant No.: GM 35654-01
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/42 (Item 42 from file: 155)
06115725 87089725
Mating type-like conversion promoted by the 2 micrograms circle site-specific recombinase: implications fo the double-strand-gap repair model.
Jayaram M
Mol Cell Biol Nov 1986, 6 (11) p3831-7, ISSN 0270-7306
Journal Code: NGY

Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/43 (Item 43 from file: 155)
06115667 87089667
Identification of the crossover site during FLP-mediated recombination in the Saccharomyces cerevisiae plasmid 2 microns circle.
McLeod M; Craft S; Broach JR
Mol Cell Biol Oct 1986, 6 (10) p3357-67, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/44 (Item 44 from file: 155)
06090546 87064546
Interaction of the FLP recombinase of the Saccharomyces cerevisiae 2 micron plasmid with mutated target sequences.
Andrews BJ; McLeod M; Broach J; Sadowski PD
Mol Cell Biol Jul 1986, 6 (7) p2482-9, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/45 (Item 45 from file: 155)
06009798 86310798
The FLP recombinase of the Saccharomyces cerevisiae 2 microns plasmid attaches covalently to DNA via a phosphotyrosyl linkage.
Gronostajski RM; Sadowski PD
Mol Cell Biol Nov 1985, 5 (11) p3274-9, ISSN 0270-7306
Journal Code: NGY
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/46 (Item 46 from file: 155)
06003314 86304314
Specific contacts between the FLP protein of the yeast 2-micron plasmid and its recombination site.
Bruckner RC; Cox MM
J Biol Chem Sep 5 1986, 261 (25) p11798-807, ISSN 0021-9258
Journal Code: HIV
Contract/Grant No.: GM32335; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/47 (Item 47 from file: 155)
05983659 86284659
Chromatin organization of the Saccharomyces cerevisiae 2 microns plasmid depends on plasmid-encoded products.
Veit BE; Fangman WL
Mol Cell Biol Sep 1985, 5 (9) p2190-6, ISSN 0270-7306
Journal Code: NGY
Contract/Grant No.: GM18926
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/48 (Item 48 from file: 155)
05980709 86281709
FLP site-specific recombinase of yeast 2-micron plasmid. Topological features of the reaction.
Beatty LG; Babineau-Clary D; Hogrefe C; Sadowski PD
J Mol Biol Apr 20 1986, 188 (4) p529-44, ISSN 0022-2836
Journal Code: J6V
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/49 (Item 49 from file: 155)
05971102 86272102
Site-specific recombination promotes plasmid amplification in yeast.
Volkert FC; Broach JR
Cell Aug 15 1986, 46 (4) p541-50, ISSN 0092-8674 Journal Code: CQ4
Contract/Grant No.: GM-34596
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/50 (Item 50 from file: 155)
05958059 86259059
The minimal duplex DNA sequence required for site-specific recombination promoted by the FLP protein of yeast in vitro.
Proteau G; Sidenberg D; Sadowski P
Nucleic Acids Res Jun 25 1986, 14 (12) p4787-802, ISSN 0305-1048
Journal Code: O8L
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/51 (Item 51 from file: 155)
05931585 86232585
Sequence organization of the circular plasmid pKD1 from the yeast Kluyveromyces drosophilarum.
Chen XJ; Saliola M; Falcone C; Bianchi MM; Fukuhara H
Nucleic Acids Res Jun 11 1986, 14 (11) p4471-81, ISSN 0305-1048
Journal Code: O8L
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/52 (Item 52 from file: 155)
05923006 86224006
Directionality in FLP protein-promoted site-specific recombination is mediated by DNA-DNA pairing.
Senecoff JF; Cox MM
J Biol Chem Jun 5 1986, 261 (16) p7380-6, ISSN 0021-9258
Journal Code: HIV
Contract/Grant No.: GM32335; AI00599
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/53 (Item 53 from file: 155)
05919123 86220123
The integrase family of site-specific recombinases: regional similarities

and global diversity.
Argos P; Landy A; Abremski K; Egan JB; Haggard-Ljungquist E; Hoess RH; Kahn ML; Kalionis B; Narayana SV; Pierson LS 3d; et al
EMBO J Feb 1986, 5 (2) p433-40, ISSN 0261-4189 Journal Code: EMB
Contract/Grant No.: AI 13544
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/54 (Item 54 from file: 155)
05810590 86111590
Site-specific recombinases: changing partners and doing the twist.
Sadowski P
J Bacteriol Feb 1986, 165 (2) p341-7, ISSN 0021-9193 Journal Code: HH3
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW

2/3/55 (Item 55 from file: 155)
05741647 86042647
The FLP recombinase of the yeast 2-micron plasmid: characterization of its recombination site.
Senecoff JF; Bruckner RC; Cox MM
Proc Natl Acad Sci U S A Nov 1985, 82 (21) p7270-4, ISSN 0027-8424 Journal Code: PV3
Contract/Grant No.: GM32335
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/56 (Item 56 from file: 155)
05707309 86008309
The FLP protein of the 2-micron plasmid of yeast. Inter- and intramolecular reactions.
Gronostajski RM; Sadowski PD
J Biol Chem Oct 5 1985, 260 (22) p12328-35, ISSN 0021-9258 Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/57 (Item 57 from file: 155)
05707308 86008308
Determination of DNA sequences essential for FLP-mediated recombination by a novel method.
Gronostajski RM; Sadowski PD
J Biol Chem Oct 5 1985, 260 (22) p12320-7, ISSN 0021-9258 Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/58 (Item 58 from file: 155)
05707307 86008307
The FLP protein of the 2-micron plasmid of yeast. Purification of the protein from Escherichia coli cells expressing the cloned FLP gene.
Babineau D; Vetter D; Andrews BJ; Gronostajski RM; Proteau GA; Beatty LG; Sadowski PD

J Biol Chem Oct 5 1985, 260 (22) p12313-9, ISSN 0021-9258
Journal Code: HIV
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/59 (Item 59 from file: 155)
05560933 85176933
The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences.
Andrews BJ; Proteau GA; Beatty LG; Sadowski PD
Cell Apr 1985, 40 (4) p795-803, ISSN 0092-8674 Journal Code: CQ4
Languages: ENGLISH
Document type: JOURNAL ARTICLE

2/3/60 (Item 1 from file: 5)
8906509 BIOSIS Number: 42131509
AN ORDERED DISASSEMBLY OF COMPLEXES OF FLP RECOMBINASE AND FRT SITES FOLLOWING RECOMBINATION
WAITE L L; COX M M
DEP. BIOCHEM., UNIV. WISCONSIN, MADISON, WIS. 53706.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 67. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/61 (Item 2 from file: 5)
8906501 BIOSIS Number: 42131501
LIGATION ACTIVITY OF THE FLP RECOMBINASE
PAN G; SADOWSKI P D
DEP. MOLECULAR MED. GENETICS, UNIV. TORONTO, TORONTO, ONTARIO M5S 1A8, CAN.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 65. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/62 (Item 3 from file: 5)
8906498 BIOSIS Number: 42131498
HALF-SITE RECOMBINATIONS MEDIATED BY FLP RECOMBINASE FROM SACCHAROMYCES-CEREVISIAE
SERRE M-C; LEI-ZHENG; JAYARAM M
DEP. MICROBIOL., UNIV. TEXAS AUSTIN, AUSTIN, TEX. 78746.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 64. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/63 (Item 4 from file: 5)
8906492 BIOSIS Number: 42131492
FUNCTIONAL ANALYSES OF MUTANTS OF FLP AND R RECOMBINASE FROM YEAST
CHEN J-W; LEE J; EVANS B; SERRE M-C; ARAKI H; OSHIMA Y; JAYARAM M

DEP. MICROBIOL., UNIV. TEXAS AUSTIN, AUSTIN, TEX. 78712.
KEYSTONE SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA REPLICATION AND RECOMBINATION, TAOS, NEW MEXICO, USA, JANUARY 25-FEBRUARY 1, 1992. J CELL BIOCHEM SUPPL 0 (16 PART B). 1992. 62. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/64 (Item 5 from file: 5)
8197568 BIOSIS Number: 91118568
TYROSINE-60 VARIANTS OF FLP RECOMBINASE GENERATE CONFORMATIONALLY ALTERED PROTEIN DNA COMPLEXES DIFFERENTIAL ACTIVITY IN FULL-SITE AND HALF RECOMBINATIONS
CHEN J-W; EVANS B R; ZHENG L; JAYARAM M
DEP. MICRBOIOL., UNIV. TEXAS AUSTIN, AUSTIN, TEX. 78712, USA.
J MOL BIOL 218 (1). 1991. 107-118. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/65 (Item 6 from file: 5)
7103760 BIOSIS Number: 88026505
FLP-FRT MEDIATED INTRACHROMOSOMAL RECOMBINATION ON A TANDEMLY DUPLICATED YE-P INTEGRANT AT THE ILV2 LOCUS OF CHROMOSOME XIII IN SACCHAROMYCES-CEREVISIAE
RANK G H; ARNDT G M; XIAO W
DEP. BIOL., UNIV. SASKATCHEWAN, SASKATOON, SASKATCHEWAN, CANADA S7N 0W0.
CURR GENET 15 (2). 1989. 107-112. CODEN: CUGED
Full Journal Title: Current Genetics
Language: ENGLISH

2/3/66 (Item 7 from file: 5)
7043154 BIOSIS Number: 87103675
FLP RECOMBINASE OF THE 2 MUM CIRCLE PLASMID OF SACCHAROMYCES-CEREVISIAE BENDS ITS DNA TARGET ISOLATION OF FLP MUTANTS DEFECTIVE IN DNA BENDING
SCHWARTZ C J E; SADOWSKI P D
DEP. MED. GENETICS, UNIV. TORONTO, TORONTO, ONTARIO M5S 1A8, CAN.
J MOL BIOL 205 (4). 1989. 647-658. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/67 (Item 8 from file: 5)
6944460 BIOSIS Number: 87004981
HIGH FREQUENCY FLP-INDEPENDENT HOMOLOGOUS DNA RECOMBINATION OF 2 MICRON PLASMID IN THE YEAST SACCHAROMYCES-CEREVISIAE
BRUSCHI C V; HOWE G A
DEP. MICROBIOL. IMMUNOL., SCH. MED., EAST CAROLINA UNIV., GREENVILLE, N.C. 27858-4354, U.S.A.
CURR GENET 14 (3). 1988. 191-200. CODEN: CUGED
Full Journal Title: Current Genetics
Language: ENGLISH

2/3/68 (Item 9 from file: 5)
6892306 BIOSIS Number: 37086685
THE FLP RECOMBINASE STEP-ARREST MUTANTS AND INTERMEDIATES IN RECOMBINATION

JAYARAM M; PARSONS R; EVANS B
RES. INST. SCRIPPS CLIN., LA JOLLA, CALIF. 92037.
SYMPOSIUM ON MOLECULAR MECHANISMS IN DNA R PLICATION AND RECOMBINATION HELD AT THE 18TH ANNUAL UCLA (UNIVERSITY OF ALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, TEAMBOAT SPRINGS, COLORADO, USA, MARCH 27-APRIL 3, 1989. J CELL BIOCHEM UPPL 0 (13 PART D). 1989. 106. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/69 (Item 10 from file: 5)
6636107 BIOSIS Number: 86102658
AUTOREGULATION OF 2-MUM CIRCLE GENE EXPRES ION PROVIDES A MODEL FOR MAINTENANCE OF STABLE PLASMID COPY LEVELS
SOM T; ARMSTRONG K A; VOLKERT F C; BROACH J R
DEP. MOLECULAR BIOL., PRINCETON UNIV., PRINCETON, NEW JERSEY 08544.
CELL 52 (1). 1988. 27-38. CODEN: CELLB
Full Journal Title: Cell
Language: ENGLISH

2/3/70 (Item 11 from file: 5)
6624830 BIOSIS Number: 86091381
THE INT FAMILY OF SITE-SPECIFIC RECOMBINASES SOME THOUGHTS ON A GENERAL REACTION MECHANISM
JAYARAM M
DEP. MOL. BIOL., RES. INST. SCRIPPS CLINIC, 10666 NORTH TORREY PINES ROAD, LA JOLLA, CALIF. 92037, USA.
J GENET 67 (1). 1988. 29-36. CODEN: JOGI A
Full Journal Title: Journal of Genetics
Language: ENGLISH

2/3/71 (Item 12 from file: 5)
6571174 BIOSIS Number: 86037725
FLP RECOMBINASE INDUCTION OF THE BREAKAGE-FUSION-BRIDGE CYCLE AND GENE CONVERSION IN SACCHAROMYCES-CEREVISIAE
RANK G H; XIAO W; KOLENOVSKY A; ARNDT G
DEP. BIOL., UNIV. SASK., SASKATOON, SASK., CAN. S7N OWO.
CURR GENET 13 (4). 1988. 273-282. CODEN: CUGED
Full Journal Title: Current Genetics
Language: ENGLISH

2/3/72 (Item 13 from file: 5)
6150196 BIOSIS Number: 35015717
PURIFICATION OF FLP RECOMBINASE USING SEQUENCE-SPECIFIC DNA AFFINITY CHROMATOGRAPHY
GATES C A; MEYER-LEON L; ATTWOOD J M; WOOD E A; COX M M
DEP. BIOCHEM., UNIV. WIS.-MADISON, MADISON, WIS. 53706, USA.
BURGESS, R. (ED.). UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY NEW SERIES, VOL. 68. PROTEIN PURIFICATION: MICRO TO MACRO; CETUS-UCLA SYMPOSIUM, FRISCO, COLORADO, USA, MARCH 29-APRIL 4, 1987. XVIII+510P. ALAN R. LISS, INC.: NEW YORK, NEW YORK, USA. ILLUS. ISBN 0-8451-2667-9. 0 (0). 1987. 197-206. CODEN: USMBD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/73 (Item 14 from file: 5)
5802738 BIOSIS Number: 83065045
SUBSTRATE RECOGNITION BY THE 2-MICROMETER CIRCLE SITE-SPECIFIC RECOMBINASE EFFECT OF MUTATIONS WITHIN THE SYMMETRY ELEMENTS OF THE MINIMAL SUBSTRATE
PRASAD P V; HORENSKY D; YOUNG L-J; JAYARAM M
DEP. MOL. BIOL., RES. INST. SCRIPPS CLIN., LA JOLLA, CALIF. 92037, USA.
MOL CELL BIOL 6 (12). 1986. 4329-4334. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/74 (Item 15 from file: 5)
5761770 BIOSIS Number: 83024077
MATING TYPE-LIKE CONVERSION PROMOTED BY THE 2 MICROMETER CIRCLE SITE-SPECIFIC RECOMBINASE IMPLICATIONS FOR THE DOUBLE-STRAND-GAP REPAIR MODEL
JAYARAM M
DEP. MOLECULAR BIOLOGY, RESEARCH INST. SCRIPPS CLINIC, LA JOLLA, CALIFORNIA 92037.
MOL CELL BIOL 6 (11). 1986. 3831-3837. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/75 (Item 16 from file: 5)
5751545 BIOSIS Number: 83013852
ASSOCIATION OF RECIPROCAL EXCHANGE WITH GENE CONVERSION BETWEEN THE REPEATED SEGMENTS OF 2-MICROMETER CIRCLE
JAYARAM M
DEPARTMENT OF MOLECULAR BIOLOGY, RESEARCH INSTITUTE OF SCRIPPS CLINIC, 10666 NORTH TORREY PINES ROAD, LA JOLLA, CALIF. 92037, USA.
J MOL BIOL 191 (3). 1986. 341-354. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/76 (Item 17 from file: 5)
5696494 BIOSIS Number: 33091515
MECHANISMS OF ACTION OF THE FLP RECOMBINASE OF THE 2-MICRON PLASMID OF YEAST
SADOWSKI P D; BEATTY L G; CLARY D; OLLERHEAD S
DEP. MED. GENETICS, MED. SCIENCES BUILD., UNIV. TORONTO, TORONTO, CANADA M5S 1A8.
MCMACKEN, R. AND T. J. KELLY (ED.). UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY NEW SERIES, VOL. 47. DNA REPLICATION AND RECOMBINATION; PARK CITY, UTAH, USA, MARCH 16-23, 1986. XXVI+782P. ALAN R. LISS, INC.: NEW YORK, NEW YORK, USA. ILLUS. ISBN 0-8451-2646-6. 0 (0). 1987. 691-702. CODEN: USMBD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/77 (Item 18 from file: 5)
5504855 BIOSIS Number: 32027162
INTERACTION OF THE FLP RECOMBINASE OF THE 2-MICRON PLASMID WITH ITS TARGET SEQUENCE

SADOWSKI P D; ANDREWS B J; BEATTY L G; SIDENBERG D; PROTEAU G
DEP. MED. GENETICS, UNIV. TORONTO, TORONTO M5S 1A8, CAN.
KLAR, A. AND J. N. STRATHERN (ED.). CURRENT COMMUNICATIONS IN MOLECULAR BIOLOGY: MECHANISMS OF YEAST RECOMBINATION; MEETING, COLD SPRING HARBOR, N.Y., USA. IX+193P. COLD SPRING HARBOR LABORATORY: COLD SPRING HARBOR, N.Y., USA. ILLUS. PAPER. ISBN 0-87969-195-6. 0 (0). 1986. 7-10. CODEN: 24607
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/78 (Item 19 from file: 5)
5426144 BIOSIS Number: 82070947
INTERACTION OF THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2 MICROMETER PLASMID WITH MUTATED TARGET SEQUENCES
ANDREWS B J; MCLEOD M; BROACH J; SADOWSKI P D
DEP. OF MED. GENETICS, UNIV. OF TORONTO, TORONTO, ONTARIO M5S 1A8, CANADA.
MOL CELL BIOL 6 (7). 1986. 2482-2489. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/79 (Item 20 from file: 5)
5389362 BIOSIS Number: 82034165
FLP SITE-SPECIFIC RECOMBINASE OF YEAST 2-MICROMETER PLASMID TOPOLOGICAL FEATURES OF THE REACTION
BEATTY L G; BABINEAU-CLARY D; HOGREFE C; SADOWSKI P D
DEP. OF MED. GENETICS, UNIV. OF TORONTO, TORONTO M5S 1A8, CANADA.
J MOL BIOL 188 (4). 1986. 529-544. CODEN: JMOBA
Full Journal Title: Journal of Molecular Biology
Language: ENGLISH

2/3/80 (Item 21 from file: 5)
5265813 BIOSIS Number: 81033120
THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2 MICROMETER PLASMID ATTACHES COVALENTLY TO DNA VIA A PHOSPHOTYROSYL LINKAGE
GRONOSTAJSKI R M; SADOWSKI P D
DEP. MED. GENET., UNIV. TORONTO, TORONTO, ONT. M5S1A8, CAN.
MOL CELL BIOL 5 (11). 1985. 3274-3279. CODEN: MCEBD
Full Journal Title: Molecular and Cellular Biology
Language: ENGLISH

2/3/81 (Item 22 from file: 5)
5256098 BIOSIS Number: 81023405
THE FLP PROTEIN OF THE 2-MICRON PLASMID OF YEAST SACCHAROMYCES-CEREVISIAE PURIFICATION OF THE PROTEIN FROM ESCHERICHIA-COLI CELLS EXPRESSING THE CLONED FLP GENE
BABINEAU D; VETTER D; ANDREWS B J; GRONOSTAJSKI R M; PROTEAU G A; BEATTY L G; SADOWSKI P D
DEP. MED. GENETICS, UNIV. TORONTO, TORONTO, M5S 1A8, CANADA.
J BIOL CHEM 260 (22). 1985. 12313-12319. CODEN: JBCHA
Full Journal Title: Journal of Biological Chemistry
Language: ENGLISH

2/3/82 (Item 23 from file: 5)

5168213 BIOSIS Number: 31057528
THE FLP RECOMBINASE OF THE 2-MICRON PLASMID OF YEAST
SADOWSKI P D; ANDREWS B J; BABINEAU-CLARY D; BEATTY L; GRONOSTAJSKI R M; PROTEAU G; SIDENBERG D
DEP. MED. GENET., UNIV. TORONTO, TORONTO M5S 1A8, CANADA.
SYMPOSIUM ON MECHANISMS OF DNA REPLICATION AND RECOMBINATION HELD AT THE 15TH ANNUAL MEETING OF THE UCLA (UNIVERSITY OF CALIFORNIA-LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, MAR. 16-23, 1986. J CELL BIOCHEM SUPPL 0 (10 PART B). 1986. 137. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/83 (Item 24 from file: 5)
4696890 BIOSIS Number: 29054205
INTERACTION OF THE FLP RECOMBINASE WITH SUBSTRATE 2-MICRON CIRCLE DNA
ANDREWS B J; BEATTY L; SADOWSKI P D
UNIV. TORONTO.
SYMPOSIUM ON YEAST CELL BIOLOGY HELD AT THE 14TH ANNUAL MEETING OF THE UCLA (UNIVERSITY OF CALIFORNIA - LOS ANGELES) SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, APR. 9-15, 1985. J CELL BIOCHEM SUPPL 0 (9 PART C). 1985. 117. CODEN: JCBSD
Language: ENGLISH
Document Type: CONFERENCE PAPER

2/3/84 (Item 1 from file: 399)
116167825 CA: 116(17)167825y PATENT
Methods for in vitro recombination of multigene families for generation of new phenotypes
INVENTOR(AUTHOR): Short, Jay M.; Sorge, Joseph A.
LOCATION: USA
ASSIGNEE: Stratagene
PATENT: PCT International ; WO 9116427 A1 DATE: 911031
APPLICATION: WO 91US2910 (910424) *US 513957 (900424)
PAGES: 204 pp. CODEN: PIXXD2 LANGUAGE: English CLASS: C12N-015/00A; C12P-019/34B; C12P-021/06B; C07H-021/00B DESIGNATED COUNTRIES: AU; CA; FI; JP; KR; NO DESIGNATED REGIONAL: AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LU ; NL; SE

2/3/85 (Item 2 from file: 399)
106208826 CA: 106(25)208826p JOURNAL
Rapid localization and characterization of random mutations within the 2.mu. circle site-specific recombinase: a general strategy for analysis of protein function
AUTHOR(S): Govind, Nadathur S.; Jayaram, Makkuni
LOCATION: Res. Inst. Scripps Clin., La Jolla, CA, 92037, USA
JOURNAL: Gene DATE: 1987 VOLUME: 51 NUMBER: 1 PAGES: 31-41 CODEN: GENED6 ISSN: 0378-1119 LANGUAGE: English

2/3/86 (Item 3 from file: 399)
104001445 CA: 104(1)1445b JOURNAL

The FLP recombinase of the yeast 2-.mu.m plasmid: characterization of its recombination site
AUTHOR(S): Senecoff, Julie F.; Bruckner, Robert C.; Cox, Michael M.
LOCATION: Coll. Agric. Life Sci., Univ. Wisconsin, Madison, WI, 53706, USA
JOURNAL: Proc. Natl. Acad. Sci. U. S. A. DATE: 1985 VOLUME: 82
NUMBER: 21 PAGES: 7270-4 CODEN: PNASA6 ISSN: 0027-8424 LANGUAGE: English

2/3/87 (Item 4 from file: 399)
102216080 CA: 102(25)216080y JOURNAL
The FLP recombinase of the 2.mu. circle DNA of yeast: interaction with its target sequences
AUTHOR(S): Andrews, Brenda J.; Proteau, Gerald A.; Beatty, Linda G.; Sadowski, Paul D.
LOCATION: Dep. Med. Genet., Univ. Toronto, Toronto, ON, Can., M5S 1A8
JOURNAL: Cell (Cambridge, Mass.) DATE: 1985 VOLUME: 40 NUMBER: 4
PAGES: 795-803 CODEN: CELLB5 ISSN: 0092-8674 LANGUAGE: English

2/3/88 (Item 1 from file: 434)
11506141 Genuine Article#: HN234 No. References: 35
Title: SITE-SPECIFIC RECOMBINATION OF 2-MU-M PLASMID OF YEAST SACCHAROMYCES-CEREVISIAE
Author(s): PUSHNOVA EA
Corporate Source: ST PETERBURG PEDIAT MED INST/ST PETERBURG//USSR/
Journal: GENETIKA, 1992, V28, N2 (FEB), P25-34
Language: RUSSIAN Document Type: ARTICLE (Abstract Available)

2/3/89 (Item 2 from file: 434)
11487805 Genuine Article#: HM053 No. References: 33
Title: SITE-SPECIFIC INTEGRATION OF THE HAEMOPHILUS-INFLUENZAE BACTERIOPHAGE HP1 - IDENTIFICATION OF THE POINTS OF RECOMBINATIONAL STRAND EXCHANGE AND THE LIMITS OF THE HOST ATTACHMENT SITE
Author(s): HAUSER MA; SCOCCA JJ
Corporate Source: JOHNS HOPKINS UNIV,SCH HYG & PUBL HLTH,DEPT BIOCHEM/BALTIMORE//MD/21205; JOHNS HOPKINS UNIV,SCH HYG & PUBL HLTH,DEPT BIOCHEM/BALTIMORE//MD/21205
Journal: JOURNAL OF BIOLOGICAL CHEMISTRY, 1992, V267, N10 (APR 5), P 6859-6864
Language: ENGLISH Document Type: ARTICLE (Abstract Available)

2/3/90 (Item 3 from file: 434)
11338662 Genuine Article#: HB304 No. References: 21
Title: EXCHANGE OF GENE ACTIVITY IN TRANSGENIC PLANTS CATALYZED BY THE CRE-LOX SITE-SPECIFIC RECOMBINATION SYSTEM
Author(s): BAYLEY CC; MORGAN M; DALE EC; OW DW
Corporate Source: USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; USDA ARS,CTR PLANT GENE EXPRESS,800 BUCHANAN ST/ALBANY//CA/94710; UNIV CALIF BERKELEY,DEPT PLANT PATHOL/BERKELEY//CA/94720

Journal: PLANT MOLECULAR BIOLOGY, 1992, V18, N2 (JAN), P353-361
Language: ENGLISH Document Type: ARTICLE (Abstract Available)

2/3/91 (Item 4 from file: 434)
11317754 Genuine Article#: GZ516 No. References: 33
Title: A FROG VIRUS-3 GENE CODES FOR A PROTEIN CONTAINING THE MOTIF CHARACTERISTIC OF THE INT FAMILY OF INTEGRASES
Author(s): ROHOZINSKI J; GOORHA R
Corporate Source: ST JUDE CHILDRENS HOSP,DEPT VIROL & MOLEC BIOL,332 N LAUDERDALE,POB 318/MEMPHIS//TN/38101; ST JUDE CHILDRENS HOSP,DEPT VIROL & MOLEC BIOL,332 N LAUDERDALE,POB 318/MEMPHIS//TN/38101
Journal: VIROLOGY, 1992, V186, N2 (FEB), P693-700
Language: ENGLISH Document Type: ARTICLE

2/3/92 (Item 5 from file: 434)
10583597 Genuine Article#: EP811 No. References: 61
Title: A NOVEL RECOMBINATOR IN YEAST BASED ON GENE-II PROTEIN FROM BACTERIOPHAGE-F1
Author(s): STRATHERN JN; WEINSTOCK KG; HIGGINS DR; MCGILL CB
Corporate Source: NCI,FREDERICK CANC RES & DEV CTR,BASIC RES PROGRAM/FREDERICK//MD/21701
Journal: GENETICS, 1991, V127, N1, P61-73
Language: ENGLISH Document Type: ARTICLE (Abstract Available)

2/3/93 (Item 6 from file: 434)
09323349 Genuine Article#: T4208 No. References: 45
Title: FLP RECOMBINASE OF THE 2-MU-M CIRCLE PLASMID OF SACCHAROMYCES-CEREVISIAE BENDS ITS DNA TARGET - ISOLATION OF FLP MUTANTS DEFECTIVE IN DNA BENDING
Author(s): SCHWARTZ CJE; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1989, V205, N4, P647-658
Language: ENGLISH Document Type: ARTICLE

2/3/94 (Item 7 from file: 434)
07863892 Genuine Article#: F8861 No. References: 37
Title: ISOLATION OF INTERMEDIATES IN THE BINDING OF THE FLP RECOMBINASE OF THE YEAST PLASMID 2-MIRON CIRCLE TO ITS TARGET SEQUENCE
Author(s): ANDREWS BJ; BEATTY LG; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1987, V193, N2, P345-358
Language: ENGLISH Document Type: ARTICLE

2/3/95 (Item 8 from file: 434)
07372665 Genuine Article#: C9356 No. References: 23
Title: INTERACTION OF THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2-MU-M PLASMID WITH MUTATED TARGET SEQUENCES
Author(s): ANDREWS BJ; MCLEOD M; BROACH J; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/; COLD SPRING HARBOR LAB/COLD SPRING HARBOR//NY/11724; PRINCETON UNIV,DEPT MOLEC BIOL/PRINCETON//NJ/08544
Journal: MOLECULAR AND CELLULAR BIOLOGY, 1986, V6, N7, P2482-2489

Language: ENGLISH Document Type: ARTICLE

2/3/96 (Item 9 from file: 434)
07260459 Genuine Article#: C1205 No. References: 44
Title: FLP SITE-SPECIFIC RECOMBINASE OF YEAST 2-MU-M PLASMID - TOPOLOGICAL FEATURES OF THE REACTION
Author(s): BEATTY LG; BABINEAUCLARY D; HOGREFE C; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: JOURNAL OF MOLECULAR BIOLOGY, 1986, V188, N4, P529-544
Language: ENGLISH Document Type: ARTICLE

2/3/97 (Item 10 from file: 434)
06806789 Genuine Article#: AUF29 No. References: 22
Title: THE FLP RECOMBINASE OF THE YEAST 2-MU-M PLASMID - CHARACTERIZATION OF ITS RECOMBINATION SITE
Author(s): SENECOFF JF; BRUCKNER RC; COX MM
Corporate Source: UNIV WISCONSIN,COLL AGR & LIFE SCI,DEPT BIOCHEM,420 HENRY MALL/MADISON//WI/53706
Journal: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 1985, V82, N21, P7270-7274
Language: ENGLISH Document Type: ARTICLE

2/3/98 (Item 11 from file: 434)
06780315 Genuine Article#: ATE60 No. References: 28
Title: THE FLP RECOMBINASE OF THE SACCHAROMYCES-CEREVISIAE 2-MU-M PLASMID ATTACHES COVALENTLY TO DNA VIA A PHOSPHOTYROSYL LINKAGE
Author(s): GRONOSTAJSKI RM; SADOWSKI PD
Corporate Source: UNIV TORONTO,DEPT MED GENET/TORONTO M5S 1A8/ONTARIO/CANADA/
Journal: MOLECULAR AND CELLULAR BIOLOGY, 1985, V5, N11, P3274-3279
Language: ENGLISH Document Type: ARTICLE

2/3/99 (Item 1 from file: 76)
1171271 82001618771
Mutations in the 2-.mu.m circle site-specific recombinase that abolish recombination without affecting substrate recognition.
Prasad, P.V.; Young, L.-J.; Jayaram, M.
Dep. Mol. Biol., Res. Inst. Scripps Clin., 10666 N. Torrey Pines Rd., La Jolla, CA 92037, USA
PROC. NATL. ACAD. SCI. USA; 84(8), pp. 2189-2193 1987
Language: English Summary Language: English

2/3/100 (Item 1 from file: 73)
8210454 EMBASE No: 91239554
Erratum: Identification of the active site tyrosine of Flp recombinase. Possible relevance of its location to the mechanism of recombination (Vol. 265 (1990) 18504-18510)
Evans B.R.; Chen J.-W.; Parsons R.L.; Bauer T.K.; Teplow D.B.; Jayaram M.
J. BIOL. CHEM. (USA), 1991, 266/11 (7312) CODEN: JBCHA ISSN: 0021-9258
LANGUAGES: English

2/3/101 (Item 2 from file: 73)

7363228 EMBASE No: 89079376
FLP recombinase of the 2 microm circle plasmid of Saccharomyces cerevisiae bends its DNA target. Isolation of FLP mutants defective in DNA bending
Schwartz C.J.E.; Sadowski P.D.
Department of Medical Genetics, University of Toronto, Toronto, Ont. M5S 1A8 Canada
J. MOL. BIOL. (United Kingdom) , 1989, 205/4 (647-658) CODEN: JMOBA ISSN: 0022-2836
LANGUAGES: English

2/3/102 (Item 1 from file: 144)
09775158 PASCAL No.: 91-0572331
Domain of a yeast site-specific recombinase (Flp) that recognizes its target site
JING-WEN CHEN; EVANS B R; SANG-HWA YANG; TEPLOW D/ B; JAYARAM M
Univ. Texas, dep. microbiology, Austin TX 78712, USA
Journal: Proceedings of the National Academy of Sciences of the United States of America, 1991, 88 (14) 5944-5948
Language: English

2/3/103 (Item 2 from file: 144)
09771721 PASCAL No.: 91-0568894
Protein-based assymmetry and protein-protein interactions in FLP recombinase-mediated site-specific recombination
XIAO-HONG QIAN; INMAN R B; COX M M
Univ. Wisconsin, coll. agricultural life sci., dep. biochemistry, Madison WI 53706, USA
Journal: Journal of biological chemistry (The), 1990, 265 (35) 21779-21788
Language: English

2/3/104 (Item 3 from file: 144)
09730857 PASCAL No.: 91-0527991
Site-specific recombination between homologous chromosomes in Drosophila
GOLIC K G
Univ. Chicago, Howard Hughes medical inst., dep; molcular genetics cell biology, Chicago IL 60637, USA
Journal: Science : (Washington, DC), 1991, 252 (5008) 958-961
Language: English

2/3/105 (Item 4 from file: 144)
09563896 PASCAL No.: 91-0354326
Tyr60 variants of Flp recombinase generate conformationally altered protein-DNA complexes : differential activity in full-site and half-site recombinations
JING-WEN CHEN; EVANS B R; LEI ZHENG; JAYARAM M
Univ. Texas at Austin, dep. microbiology, Austin TX 78712, USA
Journal: Journal of molecular biology, 1991, 218 (1) 107-118
Language: English

2/3/106 (Item 5 from file: 144)
07823248 PASCAL No.: 87-0302971
Interaction of the FLP recombinase of the saccharomyces cerevisiae 2 mu m

plasmid with mutated target sequences
NDREWS B J; MCLEOD M; BROACH J; SADOWSKI P D
Univ. Toronto, dep. medical genetics, Toronto ON M5S 1A8, Canada
Journal: Molecular and cellular biology, 1986, 6 (7) 2482-2489
Language: ENGLISH

2/3/107 (Item 1 from file: 77)
89015048 V17N02
FLP recombinase induction of the breakage-fusion-bridge cycle (BFBC) and gene conversion in Saccharomyces cerevisiae
Rank, G.H.; Xiao, W.; Kolenovsky, A.; Arndt, G.
Univ. Saskatchewan, Saskatoon, Sask., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome Poster Paper

2/3/108 (Item 2 from file: 77)
89014585 V17N02
Structure-function relationship of the sequence specific DNA binding function of the FLP recombinase
Amin, A.A.; Sadowski, P.D.
Univ. Toronto, Toronto, Ont., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome Poster Paper

2/3/109 (Item 3 from file: 77)
89014584 V17N02
FLP recombinase of 2 mu circle of S. cerevisiae bends its DNA target: An in vitro analysis
Schwartz, C.J.E.; Sadowski, P.D.
Univ. Toronto, Toronto, Ont., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome Poster Paper

2/3/110 (Item 4 from file: 77)
89013277 V17N02
Mutational analysis of the FLP site-specific recombinase of the yeast 2 micron plasmid
Sadowski, P.
Univ. Toronto, Toronto, Ont., Canada
XVIth International Congress of Genetics 8830579 Toronto (Canada) 20-27 Aug 1988
Genetics Society of Canada; National Research Council Canada; Royal

Society of Canada; Biological Council of Canada
The Proceedings of the Congress will be Published in a Special Volume of Journal Genome

2/3/111 (Item 5 from file: 77)
89012894 V17N02
Step-arrest mutants of FLP recombinase: Implications for the mechanism of recombination
Evans, B.R.; Parsons, R.; Crain, K.; Jayaram, M.
Mol. Biol. Dep., Res. Inst. Scripps Clin. and Res. Found., La Jolla, CA, USA
14th International Conference on Yeast Genetics and Molecular Biology
8830578 Espoo (Finland) 7-13 Aug 1988
European Association for Cancer Research
Subscription Department C, John Wiley & Sons Inc., 605 Third Avenue, New York, NY 10158 (USA), Abstracts will be Published in Special Issue of Journal 'Yeast' Volume 4. ISSN 0749-503X

2/3/112 (Item 1 from file: 265)
0129563 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 5R01HG00250-04 AGENCY CODE: CRISP
Directed rearrangement of the mammalian genome in vivo
PRINCIPAL INVESTIGATOR: YOUDERIAN, PHILIP A
ADDRESS: CALIF INST OF BIOLOG RESEARCH 11099 NORTH TORREY PINES ROAD LA JOLLA, CA 92037
PERFORMING ORG.: CALIFORNIA INSTITUTE OF BIOLOGICAL RES, SAN DIEGO, CALIFORNIA
SPONSORING ORG.: NATIONAL CENTER FOR HUMAN GENOME RESEARCH
FY : 92 FUNDS: $182,972

2/3/113 (Item 2 from file: 265)
0127425 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 5R01GM35654-07 AGENCY CODE: CRISP
Site specific recombination in the yeast plasmid 2 micron circle
PRINCIPAL INVESTIGATOR: JAYARAM, MAKKUNI
ADDRESS: UNIVERSITY OF TEXAS DEPT OF MICROBIOLOGY AUSTIN, TX 78712
PERFORMING ORG.: UNIVERSITY OF TEXAS AUSTIN, AUSTIN, TEXAS
SPONSORING ORG.: NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
FY : 92 FUNDS: $265,024

2/3/114 (Item 3 from file: 265)
0020434 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 9105934; 9105934 AGENCY CODE: NSF
Genetic Analysis of Pattern Formation During Drosophila Neurogenesis
PRINCIPAL INVESTIGATOR: Ellis, Hilary M Dr.
PERFORMING ORG.: Emory University, Biology, Atlanta, GA 30322
PROJECT MONITOR: Program Manager
SPONSORING ORG.: National Science Foundation, DIV OF INTEGRATIVE BIOLOGY & NEUROSCIENC, Washington, D.C., 20550
DATES: 910715 TO 920630 FY : 91 FUNDS: $69,613

2/3/115 (Item 4 from file: 265)
0019890 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 9103946; 9103946 AGENCY CODE: NSF

Generation of Mosaicism in Mice by a Site-Specific Recombinase (FLP)
PRINCIPAL INVESTIGATOR: O'Gorman, Stephen Dr.
PERFORMING ORG.: Salk Institute for Biological Studies, Gene Expression Laboratory, San Diego, CA 92128
PROJECT MONITOR: Thomas E. Brady
SPONSORING ORG.: National Science Foundation, DIV OF INTEGRATIVE BIOLOGY & NEUROSCIENC, Washington, D.C., 20550
DATES: 910315 TO 920831 FY : 91 FUNDS: $49,522

2/3/116 (Item 5 from file: 265)
0016781 DIALOG FILE NO. 265/266 FEDERAL RESEARCH IN PROGRESS
IDENTIFYING NO.: 9019220; 9019220 AGENCY CODE: NSF
Genetic Analysis in Arabidopsis
PRINCIPAL INVESTIGATOR: Signer, Ethan R Dr.
PERFORMING ORG.: Massachusetts Institute of Technology, Biology, Cambridge, MA 02139
PROJECT MONITOR: DeLill Nasser
SPONSORING ORG.: National Science Foundation, DIV OF MOLECULAR & CELLULAR BIOSCIENCES, Washington, D.C., 20550
DATES: 910201 TO 930731 FY : 91 FUNDS: $200,000

2/3/117 (Item 1 from file: 35)
01212062 ORDER NO: AADNN-59965
THE ROLE OF DNA BENDING IN FLP-MEDIATED SITE-SPECIFIC RECOMBINATION
Author: SCHWARTZ, CAROL JUDITH ELAINE
Degree: PH.D.
Year: 1990
Corporate Source/Institution: UNIVERSITY OF TORONTO (CANADA) (0779)
Source: VOLUME 52/11-B OF DISSERTATION ABSTRACTS INTERNATIONAL. PAGE 5647. 209 PAGES
ISBN: 0-315-59965-0

2/3/118 (Item 2 from file: 35)
01142876 ORDER NO: AAD90-30816
UNUSUAL DNA STRUCTURE IN SITE-SPECIFIC AND HOMOLOGOUS RECOMBINATION (RECOMBINATION)
Author: UMLAUF, SCOTT W.
Degree: PH.D.
Year: 1990
Corporate Source/Institution: THE UNIVERSITY OF WISCONSIN - MADISON (0262)
Source: VOLUME 51/09-B OF DISSERTATION ABSTRACTS INTERNATIONAL. PAGE 4199. 219 PAGES

2/3/119 (Item 3 from file: 35)
1061565 ORDER NO: AAD89-12817
ANALYSIS OF THE MAJOR DNASE I HYPERSENSITIVE SITE ON THE YEAST TWO-MICRON DNA PLASMID
Author: STRAND, ANDREW DAVID
Degree: PH.D.
Year: 1989
Corporate Source/Institution: UNIVERSITY OF MINNESOTA (0130)
Source: VOLUME 50/02-B OF DISSERTATION ABSTRACTS INTERNATIONAL. PAGE 446. 111 PAGES

2/3/120 (Item 4 from file: 35)
949308 ORDER NO: AAD87-06690
A GENETIC ANALYSIS OF FACTORS INVOLVED IN THE MAINTENANCE OF THE 2 MICRON PLASMID OF SACCHAROMYCES CEREVISIAE (CHROMATIN)
Author: VEIT, BRUCE EDWARD
Degree: PH.D.
Year: 1986
Corporate Source/Institution: UNIVERSITY OF WASHINGTON (0250)
Source: VOLUME 47/12-B OF DISSERTATION ABSTRACTS INTERNATIONAL. PAGE 4763. 97 PAGES

2/3/121 (Item 1 from file: 51)
00405585 91-03-b0028 SUBFILE: FSTA
Yeast 2 MUm vectors replicate and undergo recombination in Torulaspora delbrueckii.
Compagno, C.; Ranzi, B. M.; Martegani, E.
Correspondence (Reprint) address, B. M. Ranzi, Dipartimento di Fisiologia e Biochimica Generali, Sezione di Biochimica Comparata, Univ. di Milano, Milan, Italy
Molecular Microbiology 1989 , 3 (8) 1003-1010
LANGUAGE: English

2/3/122 (Item 1 from file: 60)
09154644
PROJ NO: NYC-186301 AGENCY : SAES NY.C
PROJ TYPE: STATE
START: 01 JUL 91 TERM: 30 JUN 92
INVEST: MACINTYRE R J
ENTOMOLOGY
CORNELL UNIVERSITY
ITHACA NEW YORK 14853

DEVELOPMENT OF A MORE EFFICIENT INSECT TRANSFORMATION SYSTEM

OBJECTIVES:The goal of the research described below is to develop a system in which DNA canbe both easily and effectively delivered to insect embryos and, using the yeast "flip recombinase" system, insure the recovery of transgenic animals at high frequencies.

PRIMARY HEADINGS: R207 Insect Control-Field Crops;A4500 Protection Against Insects;C6500 Invertebrates;F1313 Physiology-Other

2/3/123 (Item 2 from file: 60)
09091400
PROJ NO: WIS02827 AGENCY : SAES WIS
PROJ TYPE: STATE
START: 01 JUL 86 TERM: 30 NOV 96 FY: 1989
INVEST: COX M M
BIOCHEMISTRY
UNIV OF WISCONSIN
MADISON WISCONSIN 53706

THE BIOCHEMISTRY OF GENETIC RECOMBINATION

OBJECTIVES:The FLP recombinase (derived from yeast) has been purified extensively. The properties of this protein and the recombination event it catalyzes are being studied in vitro. The recombination site utilized by this protein has been defined in detail. Studies on the mechanism of action of this recombination system are now getting underway.

PRIMARY HEADINGS: R318 Noncommodity Biotechnology, Biometry; A7000 Experimental Design, Statistical Methods;C6300 Biological Cell Systems; F0114 Biochemistry and Biophysics-Other

2/3/124 (Item 1 from file: 286)
0050984 Journal Announcement: 08APR91 Doc Type: 2
Nature, 15 MAR 1991, Vol(No) 251(4999), Page(s) 1351-1355

1ST COMPANY/ORGANIZATION NAME:
Salk Institute for Biological Studies, The, USA (1921)
?

LOCUS YSCPLASM 6318 bp DNA circular PLN 31-JUL-1992
DEFINITION Yeast (S.cerevisiae) 2 micron circle plasmid, complete genome.
ACCESSION J01347 L00321 L00322 L00323 L00324 M10185 M11111 M11593 M14239
M14240 M14241 M14242 M14243 M14244 M14245 M14253 M14254 M14255
M14256 M14257 M14258 M14259 M14591 M14592 M14593 M14594 M14595
M14596 M14597 M14598 V01323
KEYWORDS DNA-binding protein; Rep-1 protein; Rep-2 protein; circular;
complete genome; d protein; plasmid; protein FLP; recombinase;
repeat region.
SOURCE Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71, p82-6B,
CV20, pMMD2, pGP20, pJFS166 (see comment).
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Endomycetales;
Saccharomycetaceae.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Hindley,J. and Phear,G.A.
TITLE Sequence of 1019 nucleotides encompassing one of the inverted
repeats from the yeast 2 micron plasmid
JOURNAL Nucleic Acids Res. 7, 361-375 (1979)
MEDLINE 80034481
REFERENCE 2 (bases 1 to 6318; 1 to 6318)
AUTHORS Hartley,J.L. and Donelson,J.E.
TITLE Nucleotide sequence of the yeast plasmid
JOURNAL Nature 286, 860-865 (1980)
MEDLINE 81012161
REFERENCE 3 (bases 3891 to 3990)
AUTHORS Broach,J.R., Guarascio,V.R. and Jayaram,M.
TITLE Recombination within the yeast plasmid 2-micron circle is
site-specific
JOURNAL Cell 29, 227-234 (1982)
MEDLINE 82259368
REFERENCE 4 (bases 3881 to 4020)
AUTHORS McLeod,M., Volkert,F. and Broach,J.R.
TITLE Components of the site-specific recombination system encoded by the
yeast plasmid 2-micron circle
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 779-787 (1984)
MEDLINE 85153059
REFERENCE 5 (bases 670 to 732)
AUTHORS Andrews,B.J., Proteau,G.A., Beatty,L.G. and Sadowski,P.D.
TITLE The FLP recombinase of the 2 micron circle DNA of yeast:
Interaction with its target sequences
JOURNAL Cell 40, 795-803 (1985)
MEDLINE 85176933
REFERENCE 6 (bases 5570 to 5605)
AUTHORS Babineau,D., Vetter,D., Andrews,B.J., Gronostajski,R.M.,
Proteau,G.A., Beatty,L.G. and Sadowski,P.D.
TITLE The FLP protein of the 2-micron plasmid of yeast: Purification of
the protein from Escherichia coli cells expressing the cloned FLP
gene
JOURNAL J. Biol. Chem. 260, 12313-12319 (1985)
MEDLINE 86008307
REFERENCE 7 (sites)
AUTHORS Gronostajski,R.M. and Sadowski,P.D.
TITLE Determination of DNA sequences essential for FLP-mediated
recombination by a novel method
JOURNAL J. Biol. Chem. 260, 12320-12327 (1985)
MEDLINE 86008308

REFERENCE 8 (sites)
AUTHORS Sutton,A. and Broach,J.R.
TITLE Signals for transcription initiation and termination in the Saccharomyces cerevisiae plasmid 2 micron circle
JOURNAL Mol. Cell. Biol. 5, 2770-2780 (1985)
MEDLINE 86284639
REFERENCE 9 (sites)
AUTHORS Gronostajski,R.M. and Sadowski,P.D.
TITLE The FLP recombinase of the Saccharomyces cerevisiae 2-micron plasmid attaches covalently to DNA via a phosphotyrosyl linkage
JOURNAL Mol. Cell. Biol. 5, 3274-3279 (1985)
MEDLINE 86310798
REFERENCE 10 (bases 667 to 739)
AUTHORS Senecoff,J.F., Bruckner,R.C. and Cox,M.M.
TITLE The FLP recombinase of the yeast 2-micron-m plasmid: Characterization of its recombination site
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82, 7270-7274 (1985)
MEDLINE 86042647
REFERENCE 11 (sites)
AUTHORS McLeod,M., Craft,S. and Broach,J.R.
TITLE Identification of the crossover site during FLP-mediated recombination in the Saccharomyces cerevisiae plasmid 2 micron circle
JOURNAL Mol. Cell. Biol. 6, 3357-3367 (1986)
MEDLINE 87089667
COMMENT [8] sites; mRNA CAP sites and poly-adenylation sites. [9] sites; FLP binding.
[7] sites; FLP cleavage.
[11] sites; FLP-mediated recombination crossover site. Draft entry and clean copy sequence for [5] kindly provided by J.Senecoff, 24-JAN-1986.
Yeast 2 micron plasmid contains two 599 bp inverted repeats separated by a large unique (UL) and a small unique (US) region. During recombination the UL and US regions invert producing two sequence forms that differ in the orientation of one unique region relative to the other. The A form is presented below. FLP is the only 2-micron circle-encoded protein needed for specific site recombination between the IRs of 2-micron circle. The minimal size of the recombination site required for efficient FLP recombinase-catalyzed recombination in vitro is no more than 28 bp, which includes parts of two 13 bp inverted repeats (positions 690-702 and 711-723) and all of an 8 bp spacer (703-710) [5]. The FLP recombinase cleaves the DNA at the boundaries of the spacer and becomes covalently linked to the spacer DNA [5],[9]. The efficiency of the recombination is reduced if the spacer in a recombinant site is increased or decreased by 1 bp, while the spacer in the second site is unaltered [5]. Recombination between two sites with identical 1-base pair additions or deletions is relatively unaffected, suggesting that pairing of sequences in the spacer regions is important in FLP-promoted recombination events [5]. The sequence asymmetry utilized by the recombinase to determine the orientation of the site is located uniquely within the spacer region. Another 13 bp direct repeat, is found at positions 676-688 [5]. FLP-mediated recombination involving two FLP sites that are inverted with respect to each other results in inversion of the DNA sequences between the sites [4]. If the participating recombination sites are in direct orientation, FLP

promotes only the excision of the intervening DNA sequences [4].
The Rep 1 and Rep proteins are involved plasmid partitioning and protein stability.
A start codon in phase with the Rep1 coding region is located at positions 1966-1964. Two CAP sites for Rep1 mRNA are located beyond the 'atg' codon (position 2008) at positions 2004 and 2005.
Complete source information:
Yeast (S.cerevisiae, strain A364A D5) DNA, clones pJDB71 [1], p82-6B [2], CV20 [3], pMMD2 [4], pGP20 [5], pJFS166 [10].

NCBI gi: 172190

```
FEATURES        Location/Qualifiers
  source        1..6318
                /organism="Saccharomyces cerevisiae"
  exon          1..545
  conflict      replace((157.160)..(157.160),"")
                /citation=[1]
  conflict      replace((289.290)..(289.290),"")
                /citation=[1]
  repeat_region 341..939
                /note="IR2"
  conflict      replace((464.466)..(464.466),"")
                /citation=[1]
  conflict      replace(558,"")
                /citation=[1]
  conflict      replace(561,"")
                /citation=[1]
  conflict      replace((622.624)..(622.624),"")
                /citation=[1]
  conflict      replace(642,"")
                /citation=[1]
  conflict      replace((665.666)..(665.666),"")
                /citation=[1]
  misc_binding  673..722
                /note="FLP recombinase binding site A [9]"
                /bound_moiety="FLP recombinase"
  conflict      replace((793.794)..(793.794),"")
                /citation=[1]
  mRNA          complement(836..2038)
                /note="Rep1 mRNA (alt.)"
  mRNA          complement(836..2017)
                /note="Rep1 mRNA (alt.)"
  mRNA          complement(836..2019)
                /note="Rep1 mRNA (alt.)"
  mRNA          complement(836..2010)
                /note="Rep1 mRNA (alt.)"
  mRNA          complement(836..2004)
                /note="Rep1 mRNA (alt.)"
  mRNA          complement(836..2005)
                /note="Rep1 mRNA (alt.)"
  CDS           complement(887..2008)
                /note="Rep 1 protein;  NCBI gi: 172192"
                /codon_start=1
                /db_xref="PID:g172192"
                /translation="MNGERLLACIKQCIMQHFQPMVYDESRCVIETTRGTFPVPDNYK
                KYKTLAFAFVGHVLNTDDTPVIEKELDWPDPALVYNTIVDRIINHPELSQFISVAFIS
                QLKATIGEGLDINVKGTLNRRGKGIRRPKGVFFRYMESPFVNTKVTAFFSYLRDYNKI
```

```
            ASEYHNNTKFILTFSCQAYWASGPNFSALKNVIRCSIIHEYISKFVEREQDKGHIGDQ
            ELPPEEDPSRELNNVQHEVNSLTEQDAEADEGLWGEIDSLCEKWQSEAEDQTEAEIIA
            DRIIGNSQRMANLKIRRTKFKSVLYHILKELIQSQGTVKVYRGSSFSHDSIKISLHYE
            EQHITAVWVYLTVKFEEHWKPVDVEVEFRCKFKERKVDG"
mRNA        2254..2841
            /note="D mRNA (alt.; 5' end +/- 3 bp)"
mRNA        2254..2861
            /note="D mRNA (alt.; 5' end +/- 3 bp)"
CDS         2271..2816
            /note="D protein;  NCBI gi: 172193"
            /codon_start=1
            /db_xref="PID:g172193"
            /translation="MPYKTAIDCIEELATQCFLSKLTDDDVSTFRRVCSKENDIIKLA
            LRIPRTIDYTSILRLLYDTLPLRSLSFNEALPLFCYSIDPAQQRQCDLRFYLRDVVKL
            ARPRKRLEMQKALLQWLPSLLSDVTLQLLNDIRIRFEEIQPNIRQTVLQIYDRTCYPS
            LNFEHPNLGVFPETDSIFEPV"
repeat_region  3714..4312
            /note="IR1"
misc_binding  3930..3979
            /note="FLP-recombinase binding site B [9]"
            /bound_moiety="FLP recombinase"
mRNA        complement(4108..5182)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5183)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5184)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5223)
            /note="REP2 mRNA (minor alt.)"
mRNA        complement(4108..5195)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5198)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5196)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5197)
            /note="REP2 mRNA (major alt.)"
mRNA        complement(4108..5212)
            /note="REP2 mRNA (minor alt.)"
CDS         complement(4308..5198)
            /note="Rep 2 protein;  NCBI gi: 172194"
            /codon_start=1
            /db_xref="PID:g172194"
            /translation="MDDIETAKNLTVKARTAYSVWDVCRLFIEMIAPDVDIDIESKRK
            SDELLFPGYVIRPMESLTTGRPYGLDSSAEDSSVSSDSSAEVILPAAKMVKERFDSIG
            NGMLSSQEASQAAIDLMLQNNKLLDNRKQLYKSIAIIIGRLPEKDKKRATEMLMRKMD
            CTQLLVPPAPTEEDVMKLVSVVTQLLTLVPPDRQAALIGDLFIPESLKDIFNSFNELA
            AENRLQQKKSELEGRTEVNHANTNEEVPSRRTRSRDTNARGAYKLQNTITEGPKAVPT
            KKRRVATRVRGRKSRNTSRV"
mRNA        join(5549..6318,1..545)
            /note="Flp mRNA"
exon        5549..6318
CDS         join(5570..6318,1..523)
            /note="recombinase (FLP);  NCBI gi: 172191"
            /codon_start=1
            /db_xref="PID:g172191"
            /translation="MPQFGILCKTPPKVLVRQFVERFERPSGEKIALCAAELTYLCWM
```

ITHNGTAIKRATFMSYNTIISNSLSFDIVNKSLQFKYKTQKATILEASLKKLIPAWEF
TIIPYYGQKHQSDITDIVSSLQLQFESSEEADKGNSHSKKMLKALLSEGESIWEITEK
ILNSFEYTSRFTKTKTLYQFLFLATFINCGRFSDIKNVDPKSFKLVQNKYLGVIIQCL
VTETKTSVSRHIYFFSARGRIDPLVYLDEFLRNSEPVLKRVNRTGNSSSNKQEYQLLK
DNLVRSYNKALKKNAPYSIFAIKNGPKSHIGRHLMTSFLSMKGLTELTNVVGNWSDKR
ASAVARTTYTHQITAIPDHYFALVSRYYAYDPISKEMIALKDETNPIEEWQHIEQLKG
SAEGSIRYPAWNGIISQEVLDYLSSYINRRI"
old_sequence replace(5583,"")
/citation=[2]
BASE COUNT 1876 a 1284 c 1179 g 1979 t
ORIGIN 1 bp upstream of EcoRI site.
1 gaattctgaa ccagtcctaa aacgagtaaa taggaccggc aattcttcaa gcaataaaca
61 ggaataccaa ttattaaaag ataacttagt cagatcgtac aataaagctt tgaagaaaaa
121 tgcgccttat tcaatctttg ctataaaaaa tggcccaaaa tctcacattg gaagacattt
181 gatgacctca tttctttcaa tgaagggcct aacggagttg actaatgttg tgggaaattg
241 gagcgataag cgtgcttctg ccgtggccag gacaacgtat actcatcaga taacagcaat
301 acctgatcac tacttcgcac tagtttctcg gtactatgca tatgatccaa tatcaaagga
361 aatgatagca ttgaaggatg agactaatcc aattgaggag tggcagcata tagaacagct
421 aaagggtagt gctgaaggaa gcatacgata ccccgcatgg aatgggataa tatcacagga
481 ggtactagac tacctttcat cctacataaa tagacgcata taagtacgca tttaagcata
541 aacacgcact atgccgttct tctcatgtat atatatatac aggcaacacg cagatatagg
601 tgcgacgtga acagtgagct gtatgtgcgc agctcgcgtt gcattttcgg aagcgctcgt
661 tttcggaaac gctttgaagt tcctattccg aagttcctat tctctagaaa gtataggaac
721 ttcagagcgc ttttgaaaac caaaagcgct ctgaagacgc actttcaaaa aaccaaaaac
781 gcaccggact gtaacgagct actaaaatat tgcgaatacc gcttccacaa acattgctca
841 aaagtatctc tttgctatat atctctgtgc tatatcccta tataacctac ccatccacct
901 ttcgctcctt gaacttgcat ctaaactcga cctctacatc aacaggcttc caatgctctt
961 caaattttac tgtcaagtag acccatacgg ctgtaatatg ctgctcttca taatgtaagc
1021 ttatctttat cgaatcgtgt gaaaaactac taccgcgata aacctttacg gttccctgag
1081 attgaattag ttcctttagt atatgataca agacactttt gaactttgta cgacgaattt
1141 tgaggttcgc catcctctgg ctatttccaa ttatcctgtc ggctattatc tccgcctcag
1201 tttgatcttc cgcttcagac tgccatttt cacataatga atctatttca ccccacaatc
1261 cttcatccgc ctccgcatct tgttccgtta aactattgac ttcatgttgt acattgttta
1321 gttcacgaga agggtcctct tcaggcggta gctcctgatc tcctatatga cctttatcct
1381 gttctctttc cacaaactta gaaatgtatt catgaattat ggagcaccta ataacattct
1441 tcaaggcgga gaagtttggg ccagatgccc aatatgcttg acatgaaaac gtgagaatga
1501 atttagtatt attgtgatat tctgaggcaa ttttattata atctcgaaga taagagaaga
1561 atgcagtgac ctttgtattg acaaatggag attccatgta tctaaaaaat acgcctttag
1621 gccttctgat acccttccc ctgcggttta gcgtgccttt tacattaata tctaaaccct
1681 ctccgatggt ggcctttaac tgactaataa atgcaaccga tataaactgt gataattctg
1741 ggtgatttat gattcgatcg acaattgtat tgtacactag tgcaggatca ggccaatcca
1801 gttctttttc aattaccggt gtgtcgtctg tattcagtac atgtccaaca aatgcaaatg
1861 ctaacgtttt gtatttctta taattgtcag gaactggaaa agtccccctt gtcgtctcga
1921 ttacacacct actttcatcg tacaccatag gttggaagtg ctgcataata cattgcttaa
1981 tacaagcaag cagtctctcg ccattcatat ttcagttatt ttccattaca gctgatgtca
2041 ttgtatatca gcgctgtaaa aatctatctg ttacagaagg ttttcgcggt ttttataaac
2101 aaaactttcg ttacgaaatc gagcaatcac cccagctgcg tatttggaaa ttcgggaaaa
2161 agtagagcaa cgcgagttgc attttttaca ccataatgca tgattaactt cgagaaggga
2221 ttaaggctaa tttcactagt atgtttcaaa aacctcaatc tgtccattga atgccttata
2281 aaacagctat agattgcata gaagagttag ctactcaatg ctttttgtca aagcttactg
2341 atgatgatgt gtctactttc aggcgggtct gtagtaagga gaatgacatt ataaagctgg
2401 cacttagaat tccacggact atagactata ctagtatact ccgtctactg tacgatacac
2461 ttccgctcag gtccttgtcc tttaacgagg ccttaccact cttttgttac tctattgatc
2521 cagctcagca aaggcagtgt gatctaagat tctatcttcg cgatgtagta aaactagcta
2581 gaccgagaaa gagactagaa atgcaaaagg cacttctaca atggctgcca tcattattat
2641 ccgatgtgac gctgcagctt ctcaatgata ttcgaatacg ctttgaggag atacagccta
2701 atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg

2761 aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat
2821 atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg
2881 catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc
2941 ttgcacttca atagcatatc tttgttaacg aagcatctgt gcttcatttt gtagaacaaa
3001 aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac
3061 agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa
3121 acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct gcatttttac
3181 agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaatctata cttctttttt
3241 gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt
3301 ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac tgtaggtccg
3361 ttaaggttag aagaaggcta ctttggtgtc tattttctct tccataaaaa aagcctgact
3421 ccacttcccg cgtttactga ttactagcga agctgcgggt gcatttttc aagataaagg
3481 catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat
3541 agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta
3601 tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact ctatgaatag
3661 ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga
3721 ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata
3781 tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg
3841 caatatttta gtagctcgtt acagtccggt gcgtttttgg ttttttgaaa gtgcgtcttc
3901 agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga gaataggaac
3961 ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag
4021 ctgcgcacat acagctcact gttcacgtcg cacctatatc tgcgtgttgc ctgtatatat
4081 atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc
4141 tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg
4201 ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc cactcctcaa
4261 ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca taccctagaa
4321 gtattacgtg attttctgcc ccttaccctc gttgctactc tccttttttt cgtgggaacc
4381 gctttagggc cctcagtgat ggtgttttgt aatttatatg ctcctcttgc atttgtgtct
4441 ctacttcttg ttcgcctgga gggaacttct tcatttgtat tagcatggtt cacttcagtc
4501 cttccttcca actcactctt tttttgctgt aaacgattct ctgccgccag ttcattgaaa
4561 ctattgaata tatcctttag agattccggg atgaataaat cacctattaa agcagcttga
4621 cgatctggtg gaactaaagt aagcaattgg gtaacgacgc ttacgagctt cataacatct
4681 tcttccgttg gagctggtgg gactaataac tgtgtacaat ccatttttct catgagcatt
4741 tcggtagctc tcttcttgtc tttctcgggc aatcttccta ttattatagc aatagatttg
4801 tatagttgct ttctattgtc taacagcttg ttattctgta gcatcaaatc tatggcagcc
4861 tgacttgctt cttgtgaaga gagcatacca tttccaatcg aatcaaacct ttccttaacc
4921 atcttcgcag caggcaaaat tacctcagca ctggagtcag aagatacgct ggaatcttct
4981 gcgctagaat caagaccata cggcctaccg gttgtgagag attccatggg ccttatgaca
5041 tatcctggaa agagtagctc atcagactta cgtttactct ctatatcaat atctacatca
5101 ggagcaatca tttcaataaa cagccgacat acatcccaga cgctataagc tgtacgtgct
5161 tttaccgtca gattcttggc tgtttcaatg tcgtccattt tggttttctt ttaccagtat
5221 tgttcgtttg ataatgtatt cttgcttatt acattataaa atctgtgcag atcacatgtc
5281 aaaacaactt tttatcacaa gatagtaccg caaaacgaac ctgcgggccg tctaaaaatt
5341 aaggaaaagc agcaaaggtg catttttaaa atatgaaatg aagataccgc agtaccaatt
5401 attttcgcag tacaaataat gcgcggccgg tgcatttttc gaaagaacgc gagacaaaca
5461 ggacaattaa agttagtttt tcgagttagc gtgtttgaat actgcaagat acaagataaa
5521 tagagtagtt gaaactagat atcaattgca cacaagatcg gcgctaagca tgccacaatt
5581 tggtatatta tgtaaaacac cacctaaggt gcttgttcgt cagtttgtgg aaaggtttga
5641 aagaccttca ggtgagaaaa tagcattatg tgctgctgaa ctaacctatt tatgttggat
5701 gattacacat aacggaacag caatcaagag agccacattc atgagctata atactatcat
5761 aagcaattcg ctgagtttcg atattgtcaa taaatcactc cagtttaaat acaagacgca
5821 aaaagcaaca attctggaag cctcattaaa gaaattgatt cctgcttggg aatttacaat
5881 tattccttac tatggacaaa aacatcaatc tgatatcact gatattgtaa gtagtttgca
5941 attacagttc gaatcatcgg aagaagcaga taagggaaat agccacagta aaaaaatgct
6001 taaagcactt ctaagtgagg gtgaaagcat ctgggagatc actgagaaaa tactaaattc
6061 gtttgagtat acttcgagat ttacaaaaac aaaaacttta taccaattcc tcttcctagc
6121 tactttcatc aattgtggaa gattcagcga tattaagaac gttgatccga aatcatttaa

6181 attagtccaa aataagtatc tgggagtaat aatccagtgt ttagtgacag agacaaagac
6241 aagcgttagt aggcacatat acttctttag cgcaaggggt aggatcgatc cacttgtata
6301 tttggatgaa tttttgag

//